Collège Doctoral du Languedoc-Roussillon

THÈSE

Pour obtenir le grade de

Docteur

Montpellier SupAgro

Délivré par le

Centre international d'études supérieures en sciences agronomiques

Montpellier

Préparée au sein de l'école doctorale SP-SA

Sciences des Procédés – Sciences des Aliments

Et de l'unité de recherche Sciences Pour l'Oenologie

Spécialité : Biochimie, chimie et technologie des aliments

Présentée par Thibaut Bontpart

VERS L'IDENTIFICATION DES MECANISMES MOLECULAIRES IMPLIQUES DANS LA GALLOYLATION DES PROANTHOCYANIDINES CHEZ LA VIGNE

Soutenue le 17 Décembre 2015 devant le jury composé de

Carsten MILKOWSKI, Chargé de Recherche, Université de Halle-Wittenberg	Rapporteur
Andrea SCHUBERT, Professeur, Université de Turin	Rapporteur
Christian CHERVIN, Professeur, Université de Toulouse	Examinateur
Nancy TERRIER, Chargée de Recherche, INRA Montpellier	Co-encadrante de thèse
Véronique CHEYNIER, Directrice de Recherche, INRA Montpellier	Directrice de thèse

RESUME

Parmi les métabolites secondaires impliqués dans la qualité du raisin et du vin, les tanins condensés ou proanthocyanidines (PAs) jouent un rôle majeur, en particulier dans l'astringence et la stabilité de la couleur du vin. Ces molécules sont également impliquées dans la défense des plantes contre des stress biotiques et abiotiques. En outre, les effets bénéfiques des PAs pour la santé humaine sont bien documentés. Les PAs de la vigne ont la particularité d'être estérifiées avec de l'acide gallique. Une réaction d'acylation appelée galloylation est responsable de cette modification. Les études montrent que la galloylation influence les propriétés œnologiques et pharmacologiques des PAs. Dans la baie de raisin, les PAs sont synthétisées dans les premiers stades de développement, principalement dans les pellicules et les pépins. Un nombre relativement faible d'étapes enzymatiques sont nécessaires pour la biosynthèse de la structure de base de ces métabolites et les gènes correspondants sont aujourd'hui largement connus chez les plantes modèles, y compris chez la vigne. Cependant, les mécanismes moléculaires impliqués dans les étapes finales, y compris la galloylation, ne sont encore que partiellement connus. Des résultats antérieurs obtenus après la recherche de Quantitaive Trait Loci (QTL) influençant la composition du raisin, et en particulier le taux de galloylation des PAs, et des études transcriptomiques après surexpression de facteurs de transcription régulant la voie de biosynthèse des PAs ont permis l'identification de gènes potentiellement impliqués dans ces étapes. Des gènes de shikimate déshydrogénase (SDH) ont été identifiés. Ces gènes interviendraient en amont, pour la biosynthèse de l'acide gallique. Trois glucosyltransférases ainsi identifiées et déjà caractérisées au laboratoire sont impliquées dans la biosynthèse de l'ester de glucose de l'acide gallique (β-glucogalline), qui servirait d'intermédiaire pour la galloylation des PAs. Ces méthodes de criblage ont également permis d'identifier 2 acyltransférases de type sérine carboxypeptidase, nommées glucose acyltransférases (GATs) qui seraient capables de catalyser la dernière étape de galloylation: le transfert de l'acide gallique depuis la β-glucogalline sur les PAs. Le premier objectif de cette thèse a été de déterminer la fonction des SDHs codées par les gènes de vigne. Certaines SDHs recombinantes produites de façon hétérologue chez *E.coli* ont la capacité de produire de l'acide gallique *in vitro*. Leur niveau d'expression au cours du développement et dans différents tissus de la baie a également été établi. Les résultats obtenus *in vitro* sont étayés par le profil métabolique (acide gallique, β-glucogalline et PAs) de hairy-roots de vigne transformées avec un gène de SDH. Le second objectif de cette thèse a été de valider la fonction des GATs par expression transitoire dans des feuilles de tabac et des tests enzymatiques *in vitro*. La transformation transitoire de feuilles de vigne avec les GATs a permis de moduler la concentration d'esters phénoliques et notamment des flavan-3-ols galloylés *in planta*. L'étude de ces gènes a été étendue aux plantes vasculaires par des analyses phylogénétiques et a permis d'identifier des motifs peptidiques potentiellement impliqués dans les mécanismes étudiés et reflétant la sub-fonctionnalisation de certains gènes. Ce travail a fourni des informations sur les bases génétiques et les mécanismes moléculaires impliqués dans la biosynthèse de l'acide gallique et son transfert en deux étapes sur les flavan-3-ols (galloylation). De nouvelles hypothèses sur l'intervention de différents transporteurs et la nature des molécules transportées pourront être formulées.

ABSTRACT

Among the secondary metabolites involved in grape berry and wine quality, condensed tannins or proanthocyanidins (PAs) play a major role, especially in astringency and color stability of wine. These molecules are also involved in plant defence against biotic and abiotic stresses. Furthermore, the beneficial effects of PAs for human health are well documented. In grapevine, PAs are partly esterified with gallic acid. An acylation reaction called galloylation is responsible for this modification. Studies show that the galloylation influences oenological and pharmacological properties of PAs. In the grape berry, PAs are synthesized in the early stages of development, mainly in skin and seeds. A relatively small number of enzymatic steps are required for the biosynthesis of the basic structure of these metabolites and the corresponding genes are now well known in model plants, including in grapevine. However, the molecular mechanisms involved in the final steps, including galloylation, are only partially known. Earlier results obtained after the search of QTL (Quantitaive Trait Loci) influencing the composition of the grape berry, especially the galloylation ratio of PAs, and transcriptomic studies after overexpression of transcription factors that regulate PAs biosynthesis pathway, have allowed the identification of genes potentially involved in these steps. Shikimate dehydrogenase (SDH) genes were identified. These genes would intervene upstream, for the biosynthesis of gallic acid. Three identified glucosyltransferases, already characterized in the laboratory, are involved in the biosynthesis of gallic acid glucose ester (β-glucogalline), which could serve as an intermediate for PAs galloylation. These screening methods have also helped to identify 2 serine carboxypeptidase-like acyltransferases, called glucose acyltransferases (GATs) which may catalyze the last step of galloylation: the transfer of gallic acid from β-glucogalline to PAs. The first objective of this thesis was to determine the function of the SDHs encoded by grapevine genes. Recombinant SDHs, produced heterologously in *E. coli*, have the capacity to generate gallic acid *in vitro*. Their level of expression during development and in different tissues of the berry was also established. *In vitro* results are supported by the metabolic profile (gallic acid, β-glucogallin and PAs) of grapevine hairy -roots transformed with a SDH gene. The second objective of this thesis was to validate the function of the GATs by transient expression in tobacco leaves and *in vitro* enzyme assays. The transient transformation of grapevine leaves with GATs allowed the modulation of the concentration of phenolic esters and notably galloylated flavan-3-ols *in planta*. The study of these genes was extended to vascular plants by phylogenetic analyses which allowed the identification of peptide motifs potentially involved in the studied mechanisms and reflecting the sub-functionalization of some genes. This work has provided information on the genetic basis and molecular mechanisms involved in the biosynthesis of gallic acid and its two-step transfer on flavan-3-ols (galloylation). New hypotheses on the intervention of different carriers and the nature of transported molecules can be proposed.

Résumé en français

INTRODUCTION

La vigne est l'une des plantes fruitières les plus répandues dans le monde. Principalement exploitée pour la production du vin, la vigne est une plante cultivée d'intérêt économique retrouvée dans de nombreuses régions tempérées du globe. Au fil des siècles, la France a su exploiter son héritage latin pour produire un vin renommé et prisé. Cependant, face à la compétition des nouveaux marchés et la baisse de la consommation de vin, la viticulture française est à un tournant de son histoire. La typicité des vins français reste cependant une valeur sûre sur le marché mondial.

Selon les données de la Food and Agriculture Organization (FAO, 2013), la France est le 5ème pays le plus producteur en raisins dans le monde, avec environ 5,5 millions de tonnes. En comparaison, le premier producteur est la Chine, avec environ 11,5 millions de tonnes, soit 15% de la production mondiale. Cependant, la France est le premier producteur de vin avec environ 46 millions d'hectolitres produits en 2014, selon l'Organisation Internationale de la Vigne et du vin (OIV).

Dans un contexte de changement climatique et face à la nécessité de réduire les intrants chimiques, la préservation de la qualité haut de gamme des vins français est un challenge. Les adaptations nécessaires pour conserver une telle qualité peuvent être réalisées au niveau des techniques de culture (p.e. contrôle de l'irrigation) ou en sélectionnant de nouveaux cépages. Dans cette optique, la connaissance des caractéristiques de la baie de raisin est d'une importance cruciale. La qualité de celle-ci est principalement caractérisée par son acidité, sa concentration en sucres, en composés phénoliques et composés d'arôme. La diversité et la quantité de ces molécules clés sont influencées par divers facteurs, notamment le cépage, le climat, les pratiques viticoles et le stade de développement. La connaissance des mécanismes qui gouvernent ces caractéristiques au cours du développement de la baie permet d'optimiser les méthodes de culture, la date de récolte et les programmes de sélection des cépages.

Parmi les composés phénoliques du raisin, les flavonoïdes sont une famille de molécules d'intérêt en œnologie, de par leur forte accumulation dans la baie et les propriétés qu'ils confèrent aux vins. Chez les plantes, ces molécules sont impliquées dans les processus de défense face aux stress biotiques et abiotiques. Les flavonoïdes ayant un impact majeur sur les propriétés organoleptiques de la baie sont les anthocyanes et les proanthocyanidines (PAs) qui sont des oligomères et polymères de flavan-3-ols, plus connues sous le nom de tanins condensés. Dans le vin rouge, les premiers influencent la couleur tandis que les autres sont responsables de l'astringence et contribuent à l'amertume et à la stabilité de la couleur. Au niveau tissulaire, la pellicule de la baie de raisin contient des anthocyanes et des PAs, tandis que les pépins sont riches en PAs. Les PAs de la vigne ont la particularité d'être estérifiées partiellement avec de l'acide gallique. La baie de raisin accumule de nombreux autres esters de composés phénoliques. Notamment, les acides hydroxycinnamiques sont généralement associés à l'acide tartrique et aux anthocyanes.

La connaissance et la maîtrise de la composition phénolique dans la baie de raisin est essentielle pour produire le style de vin recherché. La production de vin rouge nécessite une étape de macération qui va permettre d'extraire les molécules contenues dans la pellicule et/ou les pépins. D'autres aspects, en aval du processus de vinification, doivent également être pris en compte pour la qualité du produit final. En effet, la modification de la structure des molécules, l'interaction des molécules entre elles et avec d'autres molécules présentes *in vino* au cours du vieillissement peuvent influencer les propriétés organoleptiques.

Les programmes de recherche actuels concernant les composés phénoliques du raisin, et plus généralement les composés d'intérêt œnologique, ciblent l'une des étapes permettant d'aboutir au

produit final, le vin. Ainsi, la biosynthèse *in planta*, l'extraction et la modification au cours de la vinification et le devenir *in vino* des composés phénoliques sont des thèmes couramment abordés en recherche (Kennedy *et al*., 2006) et dans l'UMR Sciences Pour l'Oenologie plus particulièrement.
La valorisation des produits issus de la vinification, et notamment des composés phénoliques, attirent également l'attention des chercheurs. Par ailleurs, de nombreux projets de recherche visent à décrire l'impact des composés phénoliques (retrouvés dans le vin notamment) sur la santé humaine. En effet, leurs propriétés anti-oxydantes auraient des vertus préventives sur les maladies cardiovasculaires, les diabètes et les cancers, notamment.

Cette thèse s'inscrit dans le cadre de l'étude de la biosynthèse des composés phénoliques *in planta*.
La connaissance des mécanismes moléculaires impliqués dans la biosynthèse et le stockage des flavonoïdes au sein des cellules végétales est une approche pertinente pour contrôler la qualité des baies de raisin. Les avancées de la biologie moléculaire au cours du vingtième siècle ont permis de décrypter la voie de biosynthèse des flavonoïdes. L'étude de mutants présentant des phénotypes particuliers chez des plantes modèles a facilité l'identification des bases génétiques impliquées dans la biosynthèse des flavonoïdes. Ainsi, les enzymes responsables des étapes clés de cette voie, permettant la biosynthèse des classes majeures de flavonoïdes, et des facteurs de transcription, capables de les réguler, ont été caractérisés. Même si la vigne n'est pas classiquement considérée comme une plante modèle, la baie de raisin est un organe riche en flavonoïdes, qui se prête à l'étude des mécanismes moléculaires impliqués dans la biosynthèse, le transport et le stockage de ces molécules. De plus, le séquençage de son génome a été publié en 2007 (Jaillon *et al*., 2007), ce qui en fait la quatrième plante dont le génome a été totalement séquencé et permet aux chercheurs d'accéder à des données génétiques.

Dans les années 2000, des facteurs de transcription spécifiques régulant la biosynthèse des anthocyanes et des PAs ont été identifiés chez la vigne, et en particulier au sein du laboratoire d'accueil. Des études transcriptomiques ont permis d'établir une liste de gènes induits par des facteurs de transcription de la biosynthèse des PAs (MybPAs), et donc potentiellement impliqués dans cette voie (Terrier *et al*., 2009). De plus, une étude de génétique quantitative menée sur une population de 191 individus issus d'un croisement entre les cépages Syrah et Grenache ont mis en évidence des portions de génome (Quantitative Trait Loci, QTL) modulant la composition en PAs des raisins (Huang *et al*., 2012).
Nous nous sommes intéressés aux gènes pouvant influencer les mécanismes permettant la galloylation des flavan-3-ols, y compris la biosynthèse de l'acide gallique. En effet, la galloylation facilite l'interaction des flavan-3-ols avec les protéines, augmente l'astringence et pourrait également jouer un rôle dans les mécanismes de défense de la plante (Nimal Punyasiri *et al*., 2004). Les mécanismes moléculaires capables d'influencer le taux de galloylation des PAs *in planta* ont été étudiés chez des espèces qui accumulent des flavan-3-ols galloylés (kaki, thé). Cependant, malgré l'identification de gènes potentiellement impliqués dans ce mécanisme, aucun n'a jamais été caractérisé. La combinaison de données de transcriptomique et de génétique quantitative ont permis d'identifier un nombre restreint de gènes candidats pour la galloylation des PAs. Un travail de caractérisation fonctionnelle a été initié au cours d'une thèse précédente (thèse Fida Khater). Trois glucosyltransférases (VvGTs) induites par les facteurs de transcription MybPA ont la capacité de produire des glucose esters d'acides hydroxybenzoïques (p. e. acide gallique) et hydroxycinnamiques (p. e. acide caféique).

Le projet de ma thèse est la suite logique des expérimentations précédentes et a pour but de caractériser des gènes déjà identifiés dont la fonction n'a pu être validée, des glucose acyltransférases (GAT, de type sérine carboxypeptidase-like, SCPL). Les glucose esters produits par les VvGTs

pourraient être les substrats des GATs pour la formation d'esters de composés phénoliques: flavan-3-ols galloylés et esters tartriques d'acides hydroxycinnamiques. Les gènes annotés comme shikimate déshydrogénases (SDH), qui pourraient intervenir en amont de la voie de biosynthèse des PAs en produisant de l'acide gallique, ont également été étudiés.
Différentes stratégies de caractérisation fonctionnelle ont été mises en œuvre pour valider la fonction prédite à ces gènes.

Le premier objectif de cette thèse était d'étudier la capacité des SDHs de vigne à produire de l'acide gallique. Leur validation fonctionnelle a été menée par expression hétérologue des protéines dans la bactérie *Escherichia coli* (*E.coli*), et par surexpression dans des racines de vigne transgéniques (hairy-roots) pour l'une des SDH.
Le second objectif était de valider la fonction de galloylation des flavan-3-ols prédite pour les GATs. La capacité de ces protéines à produire des esters tartriques d'acides hydroxycinnamiques a également été étudiée. Des approches d'expression transitoire dans les feuilles d'organisme hétérologue (tabac, *Nicotiana benthamiana*) et homologue (vigne) ont été adoptées pour valider la fonction des GATs.
La capacité des enzymes produites à produire les composés phénoliques attendus a été analysée par spectrophotométrie, chromatographie liquide haute performance – voire chromatographie haute performance couplée à la spectrométrie de masse.

Cette thèse inclut 5 chapitres.
Le Chapitre 1 consiste en une étude bibliographique menée à partir de publications scientifiques. Les caractéristiques de la baie de raisin et de ses composés phénoliques principaux sont présentées. Cette partie résume les connaissances scientifiques concernant la biosynthèse, le transport et la régulation de la biosynthèse des flavonoïdes. Une attention plus particulière a été portée aux molécules étudiées: l'acide gallique et les PAs.
Le Chapitre 2 décrit le matériel et les méthodes utilisés.
Le Chapitre 3 concerne l'étude des SDHs. Une partie de ce travail est présenté sous la forme d'un manuscrit qui sera prochainement soumis. Des résultats supplémentaires non inclus dans l'article sont également présentés.
Le Chapitre 4 est consacré à l'étude des GATs. Ce chapitre inclut un manuscrit publié comme « Research review » dans New Phytologist. La seconde partie présente les travaux menés sur l'étude des GATs.
Le Chapitre 5 est une discussion générale des résultats obtenus et des perspectives envisagées pour continuer ce travail.

PRINCIPAUX RESULTATS ET DISCUSSION

Implication des shikimate déshydrogénases dans la biosynthèse d'acide gallique

La biosynthèse de l'acide gallique a souvent été débattue dans la littérature scientifique, notamment pour déterminer son précurseur (Dewick & Haslam, 1969). La phénylalanine et des métabolites de la voie du shikimate (SA), ainsi que l'acide protocatéchique, ont été proposés comme précurseurs.

Une étude de radiomarquage a démontré que l'acide gallique est synthétisé à partir de métabolites de la voie du SA dans les jeunes feuilles, mais à partir de phénylalanine dans les feuilles plus âgées (Ishikura *et al*., 1984). Le glyphosate, un herbicide qui cible une enzyme de la fin de la voie du SA, entraine une accumulation d'acide gallique, ce qui laisse penser que son précurseur est l'un des métabolites en amont de la voie (Lydon & Duke, 1988). Plus récemment, il a été démontré que plus de 90% de l'acide gallique est synthétisé par déshydrogénation du 3-déhydroshikimate (3-DHS, Werner *et al*., 2004), ce qui confirme l'hypothèse proposée dans les années 1960. Un gène, codant une enzyme capable de produire de l'acide gallique à partir de SA et de 3-DHS, a été cloné chez le noyer (Muir *et al*., 2011). La protéine codée est une SDH, une enzyme impliquée dans la voie du SA. Chez les plantes, la SDH est décrite comme une enzyme bi-fonctionnelle capable d'assurer le flux de composés carbonés dans la voie du SA menant à la biosynthèse d'acides aminés aromatiques nécessaires au métabolisme des plantes (Moore, 2004, Singh & Christendat, 2006).

Le génome de la vigne comprend 4 SDHs (*VvSDHs*). Trois gènes (nommés *VvSDH7*, *-8* et *-9*) sont regroupés les uns à la suite des autres sous la forme de cluster sur le chromosome 14. Le quatrième gène (*VvSDH5*) est isolé sur le chromosome 5.
La co-localisation de 3 gènes dont la fonction prédite est similaire pourrait résulter d'une duplication de gènes, comme cela est fréquemment reporté chez les plantes (p.e. Rizzon *et al*., 2006). Des évènements de duplication de SDH ont été rapportés pour certaines plantes ligneuses, et notamment la vigne (Tohge *et al*., 2013). Les différents gènes paralogues provenant d'un gène ancestral commun peuvent être impliqués dans la biosynthèse d'un même métabolite, avec plus ou moins d'efficacité. Dans certains cas, un ou des gènes dupliqués peuvent acquérir une nouvelle fonction et produire un métabolite différent (néo-fonctionnalisation). Un faible nombre de substitutions au niveau d'acides aminés clés pour l'activité « classique » d'une enzyme peut modifier l'affinité de substrat et la production de métabolites (p.e. Keeling *et al*., 2008).

La caractérisation fonctionnelle des VvSDHs a été abordée par expression hétérologue dans la bactérie *E.coli*. L'ADN complémentaire (ADNc, préalablement synthétisé *in vitro* à partir d'ARN) de chaque gène a été amplifié à l'aide d'amorces spécifiques à partir de péricarpe de baie de raisin (cépage Syrah) immature et cloné dans un vecteur optimisé pour l'expression hétérologue de protéines chez *E. coli*. Ce vecteur permet de fusionner la séquence clonée à un gène codant une glutathion-*S*-transférase (GST). Cette technique facilite la purification des protéines de fusion à l'aide d'une résine de glutathion sépharose. Les protéines produites chez la bactérie ont été extraites et purifiées. Les protéines de fusion ont été libérées de leur « tag » GST sous l'action de la thrombine. La pureté et la concentration des extraits protéiques ont été analysées par SDS-PAGE (Sodium dodecyl sulfate-polyacrylamide gel electrophoresis) et méthode de Bradford, respectivement. La taille de la protéine purifiée correspond au poids moléculaire attendu (~60 kDa) pour chacune des enzymes. Aucune bande correspondant à une protéine de poids moléculaire similaire n'a été observée dans la condition contrôle (vecteur vide).

Les enzymes obtenues ont permis de tester la capacité de chaque VvSDH à produire de l'acide gallique, à partir de métabolites de la voie du SA (SA et 3-DHS), en procédant à des tests enzymatiques *in vitro*. Les produits de réaction ont été injectés et analysés sur un système HPLC-DAD (High Performance Liquid Chromatography-Diode Array Detector), équipé d'une colonne en phase inverse C18, qui a permis de séparer, identifier et quantifier les molécules initiales et formées.
A partir de SA et de $NADP^+$ (Nicotinamide Adenine Dinucleotide Phosphate), à pH 9, VvSDH5 a produit la plus grande quantité d'acide gallique (~6µM). VvSDH8 et -9 ont produit des concentrations plus faibles et VvSDH7 n'a pas produit d'acide gallique.

A partir de 3-DHS, dans les même conditions, une formation spontanée d'acide gallique a été observée (~7µM) en absence d'enzyme. Seules VvSDH8 et -9 ont produit significativement plus d'acide gallique que dans la condition sans enzyme (~11 et 20µM, respectivement).
L'identité de l'acide gallique et du 3-DHS produits par réaction enzymatique ont été vérifiées par Ultra High Performance Liquid Chromatography-Diode Array Detector-Mass Spectrometry (UPLC-DAD-MS). Les valeurs obtenues sont proches de celles rapportées par Muir *et al*. (2011) concernant la production d'acide gallique dans des conditions similaires par une SDH de noyer, malgré un temps d'incubation plus court pour les VvSDHs.
Les paramètres cinétiques de chaque enzyme ont ensuite été déterminés, en suivant la variation de NADPH au spectrophotomètre.
Les résultats révèlent que VvSDH5 convertit plus rapidement le SA en 3-DHS que VvSDH8 et -9. Les 4 enzymes ont une activité shikimate déshydrogénase (à partir de SA et de $NADP^+$) maximale en conditions basiques (pH 8.5-9). Cependant, VvSDH7 présente une activité très faible par rapport aux 3 autres enzymes.Ces résultats d'activité enzymatique *in vitro* révèlent que VvSDH8 et -9 peuvent produire deux métabolites différents à partir de 3-DHS, l'acide gallique ou le SA, selon le cofacteur ($NADP^+$ ou NADPH, respectivement). L'état redox du compartiment cellulaire, et plus précisément le ratio NADPH/$NADP^+$ doit influencer l'activité catalytique de ces enzymes. Par ailleurs, la présence d'acide gallique *in planta* pourrait résulter d'une transformation non enzymatique du 3-DHS.

Le profil d'expression de chaque *VvSDH* a été déterminé par Polymerase Chain Reaction quantitative (qPCR) en utilisant le gène *Elongation Factor 1 α* (*EF1α*) comme gène de référence. Ce profil a tout d'abord été déterminé dans le péricarpe de la baie à différentes étapes de développement (9 étapes, de 11 à 99 jours après floraison). Les baies récoltées à 52 jours après floraison sont au stade véraison, étape à partir de laquelle le mûrissement de la baie débute. Toutes les *VvSDHs* présentent un pic d'expression au stade vert (35 jours après floraison) dans le péricarpe. Leur expression est relativement faible à véraison et après véraison, excepté pour *VvSDH5* dont l'expression reste élevée quelques jours après véraison.
Dans un deuxième temps, le profil d'expression relatif des *VvSDHs* a été mesuré dans des baies de raisin dont la pellicule, la pulpe et les pépins ont été séparés, et récoltées à trois stades de développement : stade vert (18 jours après floraison), véraison (52 jours après floraison) et maturité (99 jours après floraison). Les 4 gènes sont principalement exprimés au stade vert, mais dans des tissus différents : *VvSDH5* dans les pépins, *VvSDH7* dans la pulpe et les pépins, *VvSDH8* dans les pépins et *VvSDH9* dans la pellicule et les pépins. Contrairement aux autres gènes, l'expression de *VvSDH8* est quasiment spécifique des pépins immatures.

En parallèle, la quantité d'acide gallique et de son ester de glucose, la β-glucogalline (β-G), a été déterminée aux mêmes stades de développement des baies disséquées. La β-G serait le précurseur nécessaire au transfert de l'acide gallique sur les flavan-3-ols, et donc le substrat potentiel des VvGATs. La quantité de ces molécules a été analysée selon une méthode d'Ultra High Performance Liquid Chromatography couplée à un triple-quadrupole Mass Spectrometry (UHPLC-QqQ-MS, Lambert *et al*., 2015). Indépendamment du stade de développement de la baie, l'acide gallique et la β-G sont présents en plus grande quantité dans les pépins (25-500 et 10-145 $nmol.g^{-1}$ tissu, respectivement). Ces molécules sont présentes en très faible concentration dans la pulpe, parfois à la limite du seuil de détection (<1 $nmol.g^{-1}$ tissu). Des quantités intermédiaires ont été mesurées dans les pellicules (1,4-20 $nmol.g^{-1}$ tissu). La quantité maximale d'acide gallique a été mesurée dans les pépins matures (~500 $nmol.g^{-1}$ tissu), probablement liée à l'hydrolyse des flavan-3-ols galloylés. Excepté dans les pépins, la quantité d'acide gallique est la plus élevée en stade vert, suggérant que sa biosynthèse intervient avant véraison. Par ailleurs, la quantité maximale de β-G a été mesurée dans les

pépins immatures (~150 nmol.g^{-1} tissu), ce qui coïncide avec le fort taux de galloylation des flavan-3-ols au cours du stade vert dans les pépins et confirme son rôle de molécule précurseur pour la galloylation des flavan-3-ols. De plus, le fort niveau d'expression des VvGTs, des enzymes capables de produire de la β-G *in vitro* (Khater *et al*., 2012), dans les pépins immatures appuie également cette hypothèse.

Les séquences protéiques de SDH de plantes dicotylédones (100 séquences provenant de 34 espèces) et monocotylédones (31 séquences provenant de 11 espèces) disponibles sur les bases de données publiques (NCBI, Phytozome) ont été collectées. Les dicotylédones possèdent entre 1 et 6 gènes codant pour une SDH, contre 1 à 4 pour les monocotylédones. Les séquences des VvSDHs ont été ajoutées à ces données pour construire un arbre phylogénétique (Neighbor-joining tree). Les séquences de SDH de dicotylédones se répartissent dans 5 groupes distincts (I-V). VvSDH5 appartient au groupe I, VvSDH7 au groupe II, VvSDH8 au groupe III et VvSDH9 au groupe IV. Les séquences de SDH de monocotylédones se regroupent indépendamment et 4 sous-groupes ont été identifiés (M-I, M-II, M-III et M-IV). Les espèces dicotylédones dont le génome entier a été séquencé possèdent au moins un gène SDH du groupe I. Des motifs protéiques clés, permettant l'interaction de la protéine avec le substrat et le cofacteur, ont été mis en évidence chez la SDH d'*Arabidopsis thaliana* (Singh & Christendat, 2006). Un alignement de séquences a permis de comparer ces motifs. Des divergences communes ont été identifiées pour les séquences appartenant à un même groupe. Les motifs observés chez les SDH du groupe I (VvSDH5) sont similaires à ceux de la SDH d'Arabidopsis mais différents de ceux des groupes III (VvSDH8) et IV (SDH9). Les substitutions d'acides aminés dans les motifs impliqués dans la liaison et l'orientation du substrat chez VvSDH8 et -9 pourraient expliquer leur capacité à produire de l'acide gallique. Les motifs clés de VvSDH7 sont identiques à ceux de SDH ayant des fonctions de quinate déshydrogénases caractérisées chez le peuplier (Guo *et al*., 2014).
Les motifs de SDHs de plantes monocotylédones sont similaires à ceux présents chez la SDH d'Arabidopsis. Les substitutions observées chez les SDHs des groupes III et IV n'ont pas été retrouvés chez les SDH de monocotylédones. Les monocotylédones n'accumulent pas de molécules dérivées de l'acide gallique (acide ellagique, tanins hydrolysables, p.e. Bate-Smith, 1973), ce qui sous-entend qu'elles n'ont pas la capacité de produire l'acide gallique. La conservation de ces motifs chez les SDH de monocotylédones pourrait justifier leur incapacité ou leur faible capacité à produire des dérivés d'acide gallique.

L'ADNc de *VvSDH8* a également été cloné dans un vecteur de surexpression, intégré ensuite dans *Agrobacterium rhizogenes*. Des vitroplants de vigne ont été inoculés avec une solution d'agrobactéries transformées afin de générer des cals cellulaires puis des racines transgéniques appelées hairy-roots surexprimant le gène d'intérêt (*VvSDH8*). Une analyse par UHPLC-QqQ-MS des concentrations en acide gallique, β-G et flavan-3-ols a permis de comparer le profil métabolique des hairy-roots surexprimant *VvSDH8* et d'une hairy-root témoin non transformée. Les résultats révèlent que les racines transformées contiennent plus d'acide gallique (environ 20-100% de plus) et de β-G (environ 50-160% de plus). La dépolymérisation des chaines de PAs par phloroglucinolyse a permis d'identifier la structure des monomères constitutifs. Le taux de galloylation des PAs est significativement plus élevé chez les hairy-roots transformées (%G = 4,5-5 environ) comparé à la hairy-root témoin (%G= 2,7% environ). La quantité totale de flavan-3-ols est significativement différente chez une seule lignée par rapport à la lignée témoin.
Le profil métabolique de hairy-roots transgéniques a permis de confirmer la capacité de VvSDH8 à produire de l'acide gallique, observée *in vitro*. Cette enzyme a la capacité de court-circuiter le flux de composés carbonés qui transitent dans la voie du SA et d'orienter le flux vers le métabolisme de l'acide gallique, favorisant ainsi la biosynthèse de β-G et de flavan-3-ols galloylés dans les tissus.

Implication de serine carboxypeptidases-like dans la galloylation des flavan-3-ols

L'acylation des métabolites secondaires est gouvernée par deux grandes familles d'acyltransférases : les acyltransférases de type BAHD (ainsi nommées à partir de la première lettre des quatre premiers gènes caractérisées appartenant à cette famille, D'Auria *et al*., 2006) et de type SCPL, qui utilisent des esters d'acyl CoA et des esters de glucose comme molécules donneuses, respectivement (pour revue, Bontpart *et al*., 2015). Les acyltransférases de type SCPL (SCPL-ATs) auraient perdu leur fonction de protéase, ce qui témoigne de la néo-fonctionnalisation de ces enzymes. La galloylation de flavan-3-ols à partir de β-G a été démontrée à partir d'extraits enzymatiques de feuilles de thé (Liu *et al*., 2012). Cependant, le(s) gène(s) responsables de ce mécanisme n'a(ont) jamais été caractérisé(s).
Deux gènes annotés comme serine carboxypeptidases, nommés Glucose Acyltransférase (GAT) ont été identifiées comme des acteurs possibles de la galloylation des PAs. Ces gènes sont localisés sur le chromosome 3 et sont séparés d'environ 300 kb.

Les séquences protéiques codées par les gènes de vigne annotés comme sérine carboxypeptidases et celles de sérine carboxypeptidases caractérisées chez d'autres plantes ont été collectées à partir de la base de données du génome de la vigne dans le but de construire un arbre phylogénétique. Cinquante et une séquences complètes de sérine carboxypeptidase ont été identifiées dans le génome de la vigne. Les séquences se regroupent en 4 clades principaux. Douze séquences de vigne, y compris GAT1 et -2, se regroupent dans le clade IA avec toutes les séquences identifiées comme des acyltransférases dans le règne végétal. Le pourcentage d'identité de séquence élevé entre certaines VvSCPs et acyltransférases caractérisées laisse penser que ces VvSCPs pourraient catalyser une réaction d'acylation similaire ou proche chez la vigne. Huit SCPL-ATs potentielles ont été identifiées à proximité des gènes *VvGATs* sur le chromosome 3.

Trois VvGTs déjà caractérisées sont capables de produire des glucose esters de d'acides hydroxybenzoïques et hydroxycinnamiques. La co-localisation des GTs et des GATs sur le chromosome 3 laisse penser à l'existence d'un cluster fonctionnel impliqué dans le métabolisme de composés secondaires, comme cela a déjà été mis en évidence chez les plantes (Osbourn, 2010). Cependant les VvGTs et les GATs sont séparées par environ 1611 kb. Les clusters fonctionnels déjà découverts chez les plantes comportent 3 à 10 gènes et sont regroupés sur environ 35–270 kb (Nützmann & Osbourn, 2014).

Les précédentes tentatives d'expression hétérologue des GATs chez des microorganismes (*E.coli* et les levures *Saccharomyces cerevisae* et *Pichia pastoris*, thèse Fida Khater) n'avaient pas permis d'obtenir des enzymes fonctionnelles et donc de valider la fonction des gènes codants. La stratégie adoptée pour cette thèse est l'expression transitoire des GATs dans des feuilles de plante (tabac et de vigne). Chaque GAT a été clonée à partir d'ADNc de cépage Maccabeu et/ou Portan dans des vecteurs de surexpression, intégrés dans des souches d'*Agrobacterium tumefaciens*. Des feuilles de tabac (*Nicotiana benthamiana*) ont été infiltrées à l'aide d'une seringue avec une solution d'agrobactéries transformées. Environ 6 jours après inoculation, les portions de feuilles inoculées ont été récoltées. Les protéines totales ont été extraites à l'aide d'un tampon d'extraction et l'extrait enzymatique a été dessalé sur colonne avant de procéder à des tests enzymatiques *in vitro* (pH 5-6,5). La réaction de galloylation a été testée à partir de β-G et d'épicatéchine. L'analyse des produits de réaction n'a pas

permis d'observer la formation d'épicatéchine galloylée (ou épicatéchine gallate). Une seconde stratégie a été d'infiltrer directement les substrats potentiels dans les feuilles de tabac, 2 jours après agroinfiltration, pour tester la réaction de galloylation *in planta*. L'analyse des métabolites extraits de ces feuilles n'a pas permis non plus de détecter la présence d'épicatéchine gallate.

L'expression transitoire des VvGATs dans des feuilles de tabac n'a pas permis d'obtenir des enzymes fonctionnelles *in vitro*. De plus, la co-infiltration des substrats potentiels de ces enzymes (β-G et épicatéchine) dans des feuilles déjà transformées avec ces gènes n'a pas permis de reconstituer le mécanisme de galloylation dans cet organisme hétérologue. Les SCPL-ATs sont sujettes à des modifications post-traductionnelles telles que la glycosylation. Il est possible que la machinerie cellulaire du tabac ne puisse pas assurer une maturation adéquate des VvGATs de vigne. Les motifs glycosidiques greffés sur les protéines au cours de leur maturation dans le réticulum endoplasmique/ appareil de Golgi peuvent différer d'une plante à une autre (Gomord *et al*., 2010). La localisation de VvGAT1 dans des vésicules cytoplasmiques (thèse Fida Khater) suggère l'existence d'un réseau vésiculaire nécessaire au mécanisme de galloylation. Les feuilles de tabac ne produisent pas de flavan-3-ols et sont donc certainement dépourvues de ce réseau permettant la galloylation des flavan-3-ols.

Une troisième stratégie a été de réaliser des transformations transitoires de feuilles de vigne (Santos-Rosa *et al*., 2008). La composition en flavan-3-ols et en esters tartriques d'acides hydroxycinnamiques a été mesurée par HPLC. Le profil métabolique des feuilles infiltrées avec des agrobactéries portant un vecteur contrôle a été comparé à celui des feuilles infiltrées avec des agrobactéries portant un vecteur de surexpression de GAT. Deux expériences indépendantes de transformation transitoire de feuilles de vigne ont été réalisées.
Les résultats d'une première expérience révèlent que les feuilles transformées avec *VvGAT1* cloné à partir du cépage Maccabeu contiennent moins de PAs et d'esters tartriques d'acides hydroxycinnamiques mais un taux de galloylation plus élevé par rapport à la condition contrôle. Cependant, des répétitions techniques n'ont pu être réalisées par manque de matériel végétal.
Dans la seconde expérience, les feuilles transformées avec *VvGAT1* cloné à partir de Portan ont un taux de galloylation des flavan-3-ols significativement plus élevé que chez le témoin (7,6 % contre 6,3 %). Les feuilles transformées avec *VvGAT1* cloné à partir de Maccabeu ont significativement moins de flavan-3-ols mais plus d'esters tartriques d'acides hydroxycinnamiques (7,4 contre 6,1 $mg.g^{-1}$ masse fraîche). Les feuilles transformées avec GAT2 cloné à partir de Macabeu ont un taux de galloylation similaire mais une concentration d'épicatéchine gallate (0,51 contre 0,31 $mg.g^{-1}$ masse fraîche) et en esters tartriques d'acides hydroxycinnamiques (9,2 contre 6,1 $mg.g^{-1}$ masse fraîche) significativement plus élevées.
Ces résultats préliminaires sont encourageants et pourront être répétées au laboratoire pour confirmer l'effet observé sur la composition en composés phénoliques dans les feuilles de vigne. Les motifs protéiques impliqués dans la spécificité de substrat pour la molécule donneuse d'acyl n'ont pas été identifiés chez les SCPL-ATs caractérisées chez d'autres plantes. Ainsi, il est possible que VvGAT2 puisse catalyser à la fois le transfert de l'acide gallique sur les flavan-3-ols et des acides hydroxycinnamiques sur l'acide tartrique.
Chez *Arabidopsis thaliana*, une SCPL-AT a la capacité d'acyler les glucosinolates avec de l'acide benzoïque et sinapique à partir de leurs glucose ester respectifs (Lee *et al*., 2012). Ainsi, une même SCPL pourrait être impliquée dans le transfert de différents composés phénoliques (acides hydroxybenzoïques et hydroxycinnamiques) et ainsi dans la biosynthèse des différents métabolites secondaires.

PERSPECTIVES

A court terme…

Il serait intéressant de modéliser l'interaction des acides aminés clés des VvSDHs avec le substrat pour chaque groupe identifié par l'étude phylogénétique. Les divergences observées par rapport à la séquence de la SDH d'Arabidopsis pourraient influencer l'activité catalytique des SDHs. Des expériences de mutagénèse dirigée pourraient valider la fonction prédite aux membres des acides aminés caractéristiques des membres de chaque groupe. L'activité quinate déshydrogénase prédite pour VvSDH7 pourra être testée *in vitro* à partir de quinate et de NAD^+. L'impact de la surexpression des 3 autres VvSDH dans des hairy-roots de vigne pourrait confirmer le rôle prédit à chacune à partir des analyses *in silico* et *in vitro*. Malgré les différentes tentatives, une seule hairy-root transformée avec *VvSDH9* a été générée. Le profil métabolique de cette lignée sera analysé pour confirmer le rôle de *VvSDH9 in planta*. La localisation intracellulaire des VvSDHs pourra être examinée dans les hairy-roots de vigne transformées avec la construction VvSDH:GFP. Par ailleurs, il serait intéressant de transformer des plantes monocotylédones avec des VvSDHs capables de produire de l'acide gallique.

Le profil d'expression des deux GATs étant déjà déterminé (thèse Fida Khater), celui des 10 autres SCPL-ATs potentielles identifiées dans le genome de la vigne pourrait être déterminé aux différents stades de développement et dans les différents tissus de la baie de raisin. Ainsi, nous pourrions prédire si ces gènes sont impliqués dans la biosynthèse d'esters de composés phénoliques au stade vert (flavan-3-ols galloylés, esters tartriques d'acides hydroxycinnamiques) ou plus tard à maturité (anthocyanes acylées). Des travaux de validation fonctionnelle pourraient également être entrepris pour ces gènes candidats.
L'expérience de transformation transitoire de feuilles de vigne devra être répétée pour confirmer les résultats obtenus. La présence de β-G in planta peut être un facteur limitant pour la galloylation des flavan-3-ols. La co-infiltration de ce substrat avec la solution d'agrobactéries permettrait d'augmenter la disponiblité de ce substrat potentiel des VvGATs *in planta*.

Sur le plus long terme…

La validation *in planta* des gènes étudiés pourrait être envisagée par la transformation stable de vitroplants de vigne, stratégie non envisageable dans le cadre d'une thèse de 3 ans.
L'hypothèse d'une galloylation des flavan-3-ols en deux étapes doit impliquer plusieurs gènes codant des transporteurs de précurseurs (d'acide gallique, β-G, flavan-3-ols), leur permettant de traverser les membranes cellulaires. Ces transporteurs n'ont jamais été identifiés chez les plantes et restent à identifier.
D'autres acteurs moléculaires pourraient être spécifiquement impliqués dans la régulation de la galloylation des flavan-3-ols. La poursuite de l'exploitation des données de transcriptomique et une étude de genome-wide association (GWA) à partir d'une core-collection de cépages de vigne permettraient de mettre en évidence de nouveaux gènes candidats pour la galloylation des PAs, mais également pour la biosynthèse d'autres esters phénoliques tels que les esters tartriques d'acides hydroxycinnamiques et les anthocyanes acylées.
Les gènes de VvSDHs et VvGATs pourraient être utilisés pour le développement de marqueurs génétiques relatifs au taux de galloylation des flavan-3-ols dans les programmes de création de

nouveaux cépages, mieux adaptés aux contraintes de la viticulture telles que la lutte contre les pathogènes et le maintien de qualité de la baie.
L'identification de gènes orthologues chez les dicotylédones ouvre la voie pour la caractérisation biochimique de SDHs impliqués dans la biosynthèse de l'acide gallique. Les mécanismes décrits pour la galloylation des flavan-3-ols chez la vigne pourraient être transposés à des plantes qui produisent ces métabolites (kaki, thé), des tanins hydrolysables ou d'autres dérivés d'acide gallique.

CONCLUSION

Les connaissances acquises au sein de l'UMR Science pour l'oenologie concernant les composés phénoliques de la baie de raisin et les techniques d'analyse mises à disposition (notamment la Plate-forme d'analyse des polyphénols) ont permis d'apporter des éléments de réponse à la question scientifique initiale. Ce travail a permis de cumuler des informations concernant les bases génétiques et les mécanismes impliqués dans la biosynthèse de l'acide gallique et des flavan-3-ols galloylés.

Remerciements

Je remercie Véronique Cheynier d'avoir accepté d'être ma directrice de thèse et pour ses remarques tout au long de mes travaux en cours.
Je tiens à remercier Mr Carsten Milkowski et Mr Andrea Schubert d'avoir accepté d'être rapporteur de mon travail, et Mr Christian Chervin d'avoir accepté de faire partie du jury.
Merci aux membres de mon comité de thèse, Alain Hehn, Isabelle Debeaujon et Agnès Ageorges, de m'avoir aidé à faire avancer mon projet de thèse.
Je remercie tout particulièrement Nancy Terrier de m'avoir encadré au cours de ces 3 années. Tes conseils et tes intuitions m'ont permis d'orienter efficacement mes travaux de recherche et de rédaction. Merci pour ta disponibilité et ta réactivité.
Merci aux membres de l'équipe qui m'ont aidé et soutenu pendant les manips': Sandrine Vialet pour la bio mol' et Jean-Luc Guiraud pour les analyses HPLC dans un premier temps, puis Thérèse Marlin qui a pris le relais et également Jean-Paul Mazauric.
Merci aux membres de la Plate-forme polyphénols qui ont participé à mes travaux et tout particulièrement Emmanuelle Meudec et Lucie Pinasseau pour les analyses en spectrométrie de masse.
Mes remerciements vont également aux membres du labo qui m'ont entouré au cours de ces 3 années de thèse, je pense notamment à Martine Pradal, Pascale Sarni-Manchado et Agnès Ageorges avec qui j'ai pu échanger au cours des pauses café matinales.
Merci à Carine Alcon de m'avoir initié à la culture et à la transformation des plants de tabac et à Pere Mestre de m'avoir accueilli à l'INRA de Colmar et transmis son savoir concernant la transformation de feuilles de vigne.
Merci aux membres de l'équipe pédagogique de l'Université de Montpellier qui m'ont permis de mener à bien ma mission d'enseignement au cours de ces 3 années, en parallèle de mes travaux de thèse, notamment Laurence Marquès et Bruno Touraine.
Merci à toutes les personnes qui ont participé de près ou de loin à la réalisation et à l'aboutissement de cette thèse.

Merci à ma famille et mes amis pour leur soutien et leurs encouragements.

Summary

Abbreviations list

3-DHS	3-Dehydroshikimate
4CL	4-Coumaroyl:CoA-Ligase
ABA	Abscisic Acid
ABC	ATP Binding Cassette
ACP	Anthocyanoplastes
ADT	Arogenate Dehydratase
AM	AnthoMATE
ANR	Anthocyanidin Reductase
ANS	Anthocyanidin Synthase
ATP	Adenosine Triphosphate
AVIs	Anthocyanic Vacuolar inclusions
BAN	BANYULS
bHLH	Basic Helix-Loop-Helix
bp	base pairs
BTL	Bilitranslocase
BZ2	Bronze 2
C3H	*p*-Coumarate 3-hydroxylase
C4H	Cinnamate-4-Hydroxylase
cDNA	Complementary Deoxyribonucleic Acid
CHI	Chalcone Isomerase
CHS	Chalcone Synthase
Ci	Initial Concentration
COMT	Caffeic acid *O*-Methyltransferase
CM	Chorismate Mutase
Cm	Centimeter
CS	Chorismate Synthase
Ct	Cycle threshold
cv.	Cultivar
DAD	Diode Array Detector
Daf	Days after flowering
DAHP	3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate
DAHPS	3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase
DFR	Dihydroflavonol 4-Reductase
DHQS	3-Dehydroquinate Synthase
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleoside 5'Triphosphate
DQD	3-Dehydroquinate Dehydratase
DW	Dry Weight
E4P	D-Erythrose 4-Phosphate
EC	Epicatechin
ECG	Epicatechin 3-*O*-Gallate
EDTA	Ethylene diamine tetra acetic Acid
EF1α	Elongation Factor 1 α
EGC	Epigallocatechin
EPSP	5-*enol*pyruvylshikimate-3-phosphate
EPSPS	5-*enol*pyruvylshikimate-3-phosphate synthase
ESI	Electrospray Ionization
EST	Expressed Sequence Tag
F3'5'H	Flavonoid 3,5'-Hydroxylase
F3'H	Flavonoid 3'-Hydroxylase
F3H	Flavanone 3-hydroxylase
FAO	Food and Agriculture Organization
FLS	Flavonol Synthase
FMN	Flavin Mononucleotide
FW	Fresh Weight
GA	Gallic Acid
GAT	Glucose Acyltransferase
GC	Gallocatechin
GFP	Green Fluorescent Protein
GRP	Grape Reaction Product
GST	Glutathion S-Transferase
GT	Glucosyltransferase
HBA	Hydroxybenzoic Acid

HCA	Hydroxycinnamic Acid
HCALDH	Hydroxycinnamaldehyde Dehydrogenase
HF	High Fidelity
HPLC	High Performance Liquid Chromatography
HT	Hexose Transporter
IAA	Indole-3-Acetic Acid
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
LAR	Leucoanthocyanidin Reductase
LB	Luria Berthani
LDOX	Leucoanthocyanin Dioxygenase
MATE	Multidrug And Toxic compound Extrusion
mDP	Mean Polymerization Degree
MES	2-(*N*-morpholino)ethane sulfonic acid
MGL	Mannitol Glutamate Luria
MRM	Multiple Reaction Monitoring
MRP	Multidrug Resistance associated Protein-carrier
MS	Mass Spectrometry
MS/2	Half-strength Murashige & Skoog
Myb	V-myb Myeloblastosis viral oncogene homolog
NADPH	Nicotinamide Adenine Dinucleotide Phosphate Hydrogen
OD	Optical Density
OIV	International Organization of Vine and Wine
ORF	Open Reading Frame
PA	Proanthocyanidin
PAL	Phenylalanine Ammonia Lyase
PAP1	Production of Anthocyanin Pigment 1
PAT	Prephenate Aminotransferase
PBS	Phosphate Buffer Saline
PCA	Protocatechuic Acid
PCR	Polymerase Chain Reaction
PDT	Prephenate Dehydratase
PEP	Phophoenolpyruvate
pH	Hydrogen Potential
PPase	Pyrophosphatase
PPO	Polyphenol Oxidase
PPY-AT	Phenylpyruvate Aminotransferase
PVPP	Polyvinylpolypyrrolidone
qPCR	quantitative Polymerase Chain Reaction
QqQ	Triple Quadrupole
qs	*Quantum Satis*
QTL	Quantitative Trait Loci
ref	*reduced epidermal fluorescence*
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
S3P	Shikimate 3-Phosphate
SA	Shikimic Acid
SCPL	Serine Carboxypeptidase Like
SDH	Shikimate Dehydrogenase
SDS	Sodium Dodecyl Sulfate
SK	Shikimate Kinase
SOC	Super Optimal broth with Catabolite repression
sp.	Species
STS	Stilbene Synthase
TAE	Tris-acetate-EDTA
Taq	*Thermus aquaticus*
TE	Tris-EDTA
TF	Transcriptional Factor
TFA	Trifluoro acetic Acid

tt	transparent testa
UFGT	UDP-glucose: Flavonoid 3-*O*-Glucosyl Transferase
UK	United Kingdom
UPLC	Ultra Performance Liquid Chromatography
UV	Ultra-Violet
Vi	Initial Velocity
V_{max}	Maximum Velocity
Vr	Reaction Volume
Vv	*Vitis vinifera*
WDR	Tryptophan-aspartic acid (W-D) Dipeptide Repeat
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
β-G	β-Glucogallin
%G	Galloylation rate

Units

°C	Celsius degree
µF	Microfarad
µg	Microgram
µL	Microliter
µm	Micrometer
µM	Micromolar
Ω	Ohm
kDa	Kilodalton
kg	Kilogram
kV	Kilovolt
L	Liter
mg	Milligram
mL	Milliliter
nm	Nanometer
Rpm	Rotation per minute
V	Volt

FIGURES & TABLES

Annexes

GENERAL INTRODUCTION

Grapevine is one of the most widespread fruit plant in the world. Mainly cultivated for winemaking, the grapevine is a crop of economic interest found in numerous world regions with temperate climate. Over the centuries, France has exploited its Latin heritage to produce a renowned and prized wine. However, faced with competition from new markets and the decline in consumption of this beverage, the French wine industry is at a crossroads in its history. The typical characteristics of French wines yet remains a safe bet in the global market.

According to FAO (Food and Agriculture Organization) data, France was the 5th largest producer of grapes in the world with about 5.5 million tonnes in 2013. By comparison, the first producer was China with about 11.5 million tons, or 15% of world production. However, France is the first world producer of wine with about 46 million hectoliters produced in 2014, according to the OIV (International Organization of Vine and Wine).

In a context of climate change and consumer claim for decrease of chemical inputs, preserving the high quality range of French wines is a challenge. The adaptations for the conservation of such quality can be reached by cultivation techniques (e.g. water status control) or selection of new cultivars. In this scope, knowledge of grape berry characteristics remains of paramount importance. The quality of a grape berry is mainly characterized by its acidity and the concentration of sugars, phenolic and aroma compounds. The study of the mechanisms that govern these characteristics during development of the berry is essential to optimize cultivation, choice of harvest date and selection process.

The diversity and content of the berry in its different metabolites are influenced by various factors, notably the grape cultivar, climate, cultural practices, development stages of the fruit.

Among the grape phenolic compounds, flavonoids are a class of interest in oenology, accumulating in high proportion in the berry and conferring organoleptic properties. In plants, these molecules are involved in defence processes against (a)biotic stress. The flavonoids which have a major impact on the organoleptic qualities of the berry are on the one hand anthocyanins, and on the other hand flavan-3-ols, found as monomers and polymers. Flavan-3-ols oligomers and polymers are called proanthocyanidins (PAs), better known as condensed tannins. In red wine, anthocyanins influence the colour while PAs are responsible for astringency, bitterness and stability of the colour. At the tissue level, the grape berry skin contains anthocyanins and PAs, while the seeds are rich in PAs. Grapevine PAs have the particularity to be acylated with gallic acid (GA) on certain subunits. This substitution is called galloylation and influences oenological and taste properties of flavan-3-ols. Grape berry also accumulate numerous esters of phenolic compounds. In particular, hydroxycinnamic acids (HCAs) can be associated to tartaric acid and anthocyanins.

The distribution of phenolic compounds in the grape berry is taken advantage for elaboration of different wine styles. The production of red wine includes a maceration step that extracts the molecules of the skin and/or seeds. Other aspects, downstream of the wine production chain, are also to be taken into account for the final product quality. Indeed, the modification of the structure, the interaction of molecules between them and with others *in vino* during ageing may affect their organoleptic properties.

Ongoing research programs on the phenolic compounds of grape berry, and more generally on the compounds of oenological interest, usually target one of the steps leading to the final product, wine. Thus, *in planta* biosynthesis, extraction and modification during the wine making process and *in vino* fate of phenolic compounds are themes commonly discussed in research (Kennedy *et al*., 2006). Valorisation of wine industry by-products, including phenolic compounds is also of interest for research. Otherwise, phenolic compounds found in wine have beneficial properties for human health. Indeed, the antioxidant properties of these molecules have therapeutic properties and many research

projects are intended to describe the impact of these compounds on cardiovascular diseases, diabetes and cancer among others.

This thesis fits into the framework of the study of the *in planta* biosynthesis of phenolic compounds.

Understanding the molecular mechanisms involved in the biosynthesis of flavonoids and storage across the cell is a relevant approach in controlling the quality of the grape berries. Advances in molecular biology during the late 20th century have made it possible to decipher the flavonoid biosynthetic pathway. The study of particular mutant phenotypes in model plants has facilitated the identification of the genetic bases involved in flavonoid biosynthesis. Thus, the enzymes responsible for the key steps in this pathway, allowing the biosynthesis of major classes of flavonoids, and transcription factors (TFs) able to regulate them were characterized.

Even if grapevine is not classically considered as a model plant, the grape berry is a flavonoid rich organ, suitable for the study of molecular mechanisms of biosynthesis, transport and storage of these molecules. Moreover, its genome has been released in 2007 (Jaillon *et al*., 2007), being therefore the 4th plant whose genome has been fully sequenced. In the 2000s, specific TFs involved in anthocyanins and PAs biosynthesis have been identified in grape, and particularly within the host laboratory.

Transcriptomic studies have established a list of genes induced by specific TFs of the biosynthesis pathway of PAs (MybPAs), and thus potentially involved in PA biosynthesis (Terrier *et al*., 2009a). Moreover, a quantitative genetic study on a population of 191 individuals from a cross between Syrah and Grenache highlighted portions of the genome (Quantitative Trait Loci, QTL) affecting the PA composition of the grape (Huang *et al*., 2012). Cross-data mining led to the identification of candidate genes for the biosynthesis of PAs.

We focused on the mechanisms leading to the galloylation of flavan-3-ols, including the biosynthesis of GA. A functional validation work of candidate genes was initiated during a previous PhD. The project of my PhD work is therefore the logical continuation of this earlier investigation and aims to characterize genes already identified but whose function could not be validated, glucose acyltransferases (also called serine carboxypeptidase-like acyltransferases), and 4 genes annotated as shikimate dehydrogenases.

The functional validation of shikimate dehydrogenases was conducted by heterologous expression of the encoded proteins in the bacterium *Escherichia coli*. Approaches of transient expression in heterologous (tobacco) or homologous (grapevine) organisms were undertaken to attempt to validate the function of glucose acyltransferases. The ability of these enzymes to synthesize some phenolic compounds was analysed by spectrophotometry, high performance liquid chromatography (HPLC) coupled to mass spectrometry – nay to diode array detection (DAD) and mass spectrometry (MS).

This manuscript includes 5 chapters.

Chapter 1 is a summary of the literature study. The characteristics of the grape berry, its major phenolic compounds are presented. This section also summarizes the scientific knowledge on the biosynthesis, transport and regulation of the production of flavonoids. More attention was paid to the molecules studied: GA, and flavan-3-ols.

Chapter 2 is the description of the material and methods.

Chapter 3 concerns the study of shikimate dehydrogenases (SDHs). A part of the work is presented in the form of a manuscript that will be submitted. Additional results not included in this article are also presented.

Chapter 4 concerns the study of glucose acyltransferases (GATs). This chapter includes a manuscript published as a "Research review" in New Phytologist. The second part presents the work performed on GATs.

Chapter 5 is a general discussion of the results obtained and the prospects envisaged to continue this work.

This work has been funded by a PhD fellowship from the French Ministère de l'Enseignement surpérieur et de la Recherche from 01/10/12 to 30/09/15.
This PhD fellowship has been attributed through the competition of Sciences des Procédés - Sciences des Aliments graduate school of Montpellier.

Research works publication

Manuscript published in a peer-reviewed journal:

- **Bontpart T, Cheynier V, Ageorges A, Terrier N. 2015.** BAHD or SCPL acyltransferase? What a dilemma for acylation in the world of plant phenolic compounds. New Phytologist **208**: 695–707.

In preparation :

- **Bontpart T, Marlin T, Vialet S, Guiraud JL, Pinasseau L, Meudec E, Cheynier V, Terrier N.** Grapevine shikimate dehydrogenases involvement in gallic acid biosynthesis. To be submitted (Journal of Experimental Botany).
- Valorisation of works concerning Glucose acyltransferases after additional experiments.

International meetings

Oral presentation:
Bontpart T, Marlin T, Vialet S, Guiraud JL, Meudec E, Ageorges A, Cheynier V, Terrier N. 2014. Grape shikimate dehydrogenases involvement in gallic acid biosynthesis. 3rd Annual Conference of the COST FA1106 QualityFruit Sept 2014, Chania (Greece).
Poster:
Bontpart T, Marlin T, Vialet S, Guiraud JL, Meudec E, Ageorges A, Cheynier V, Terrier N. 2015. Biosynthesis of gallic acid in grapevine by shikimate dehydrogenases. SEB Annual Meeting 2015, 30th June – 3rd July, Prague (Czech Republic).

Teaching activity

- 3 years of teaching at Montpellier University (about 64 hours per year, mainly in pratical work and tutorial classes) in plant physiology, phytochemicals, plant genetic engineering
- Part-time teacher at Montpellier University (October 2015-August 2016)

CHAPTER 1: BIBLIOGRAPHY

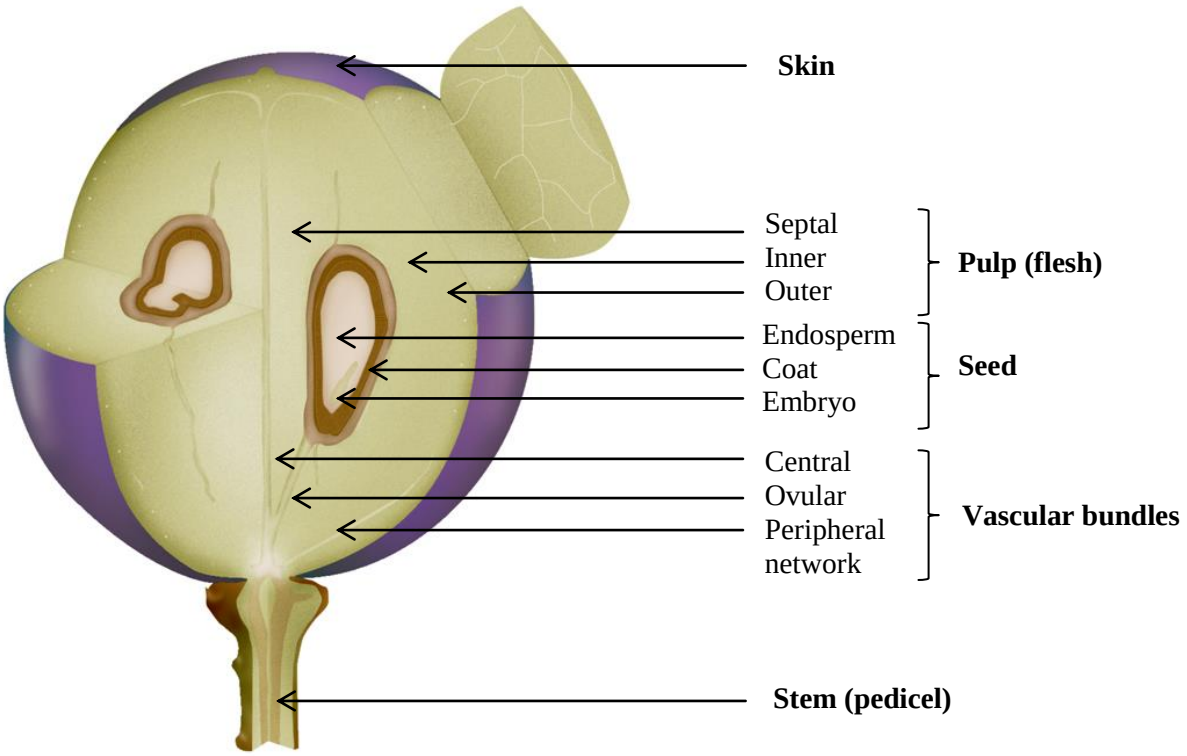


Figure 1. Ripe grape berry sectioned on the long and central axis to show internal parts.
Illustration by Jordan Koutroumanidis, Winetitles.

1. The grape berry

1.1. Morphology

The fruit of the grapevine (*Vitis vinifera* L.), called grape, is a fleshy berry composed of 3 main parts: the skin, the pulp (or flesh) and the seeds (Fig. 1). The pericarp includes the skin (exocarp), the pulp (mesocarp) and the endocarp which surround the seeds (Hardie *et al*., 1996).

The skin represents 10 to 15% of the berry weight and is subdivided into a cuticle, an epiderm and a hypoderm, from the outside to the inside. The cuticle is covered with a waxy layer which ensures water tightness to the skin and prevents the loss of water by evaporation. The epiderm consists of a single layer of little isodiametric cells. The hypoderm includes about 10 cells layers constituted of larger cells containing pigments and odorous molecules. The exact number of hypodermic layers depends on the cultivars (Hardie *et al*., 1996).

The pulp is composed of 25-30 cell layers (Hardie *et al*., 1996). Mesocarp cells are isodiametric during the green stage and expand at *véraison*. This tissue is divided in outer and inner mesocarps, separated by vascular tissues. Outer mesocarp cells can reach 500 µm whereas inner mesocarps cells do not exceed 100 µm (Hardie *et al*., 1996). The vascular tissues consist of xylem, phloem and vascular parenchyma.

The seeds are composed of a cuticle, an epidermis, an external integument (large parenchymal cells), a medium integument (2 cell layers), and an inner integument (3 cell layers) which surrounds the endosperm and the embryo (Cadot *et al*., 2006). The seeds represent at most 6% of the berry weight and contain lipids, proteins and phenolic compounds (PAs).
Lipids represent 10-16% of seed weight (Ohnishi *et al*., 1990). Values reported in different studies for protein content vary from 8% (Igartuburu *et al*., 1991) to 26% (Fazio *et al*., 1983) of the seed components.

1.2. Development

The development of the grape berries found in a same grape bunch is heterogeneous (Coombe & McCarthy, 2000).
The number of seeds per berry depends on environmental conditions at flowering which influence fertilization (Ebadi *et al*., 1995). After fertilization, the ovary wall of the carpel (pericarp) develops and becomes almost completely fleshy at maturity. The number of seeds in a berry ranges between 1 and 4 and is positively correlated to the final berry size (Hardie & Aggenbach, 1996, May, 2000).

The berry follows a typical double sigmoid curve that can be separated in 3 phases (Coombe & Hale, 1973, Fig. 2). During the development of the berry, we can observe a change in size, colour, and

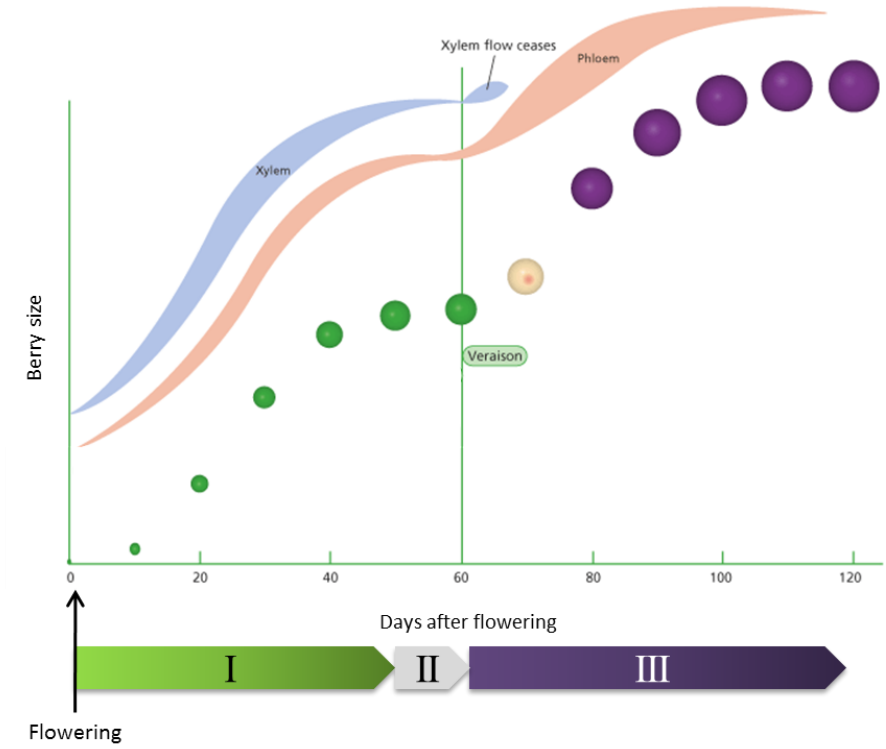


Figure 2. Phases of grape berry development according to Coombe & Hale, 1973.
I: Herbaceous phase. II: Lag phase. III: Maturation phase. Berry size and colour is modelled at 10 days intervals along the 120 days of development from flowering. Xylem flow (blue) supplies the berry during phase I and stops at *véraison* (60 days after flowering). Phloem flow (red) supplies the berry during phase III. Adapted from an illustration by Jordan Koutroumanidis, Winetitles.

storage of compounds. The water flux is the driving force of berry growth. The berry must benefit of a gradient in total water potential with the rest of the plant (Matthews & Shackel, 2005).
The duration of each phase and the concentration of compounds of oenological importance can be influenced by biotic (pathogens) and abiotic (temperature, water status and nitrogen, brightness) interactions of the plant with its environment, and also depends on the variety.

Phase I: Herbaceous phase
The first phase of growth, which starts at flowering and ends about 60 days later, is the herbaceous phase (Phase I). At this stage, the berry is little, hard and green. Cells divide and then expand rapidly (Ojeda *et al*., 1999), seeds embryos are formed during this phase. Pericarp final cells number is established during phase I (Harris *et al*., 1968). The xylem sap is the main source of water during phase I (Greenspan *et al*., 1994). It imports the ascending sap providing water and nutrients from the roots, necessary for the berry development during phase I.

The main organic acids (tartaric and malic acids) involved in the berry acidity are synthesized and accumulated during this phase. At the cellular level, these compounds are stored in the vacuole where the pH remains very acidic throughout this period (pH = 2.5, Ruffner, 1982). Hydroxycinnamic acids (HCAs) and PAs are synthesized during phase I too. Other molecules such as minerals, amino acids and flavour compounds are also accumulated during this period and contribute to the quality of the berry.

Phase II: Lag phase
The end of phase I is marked by a stationary phase (Phase II), lasting from 8 to 12 days, during which the berry is no longer growing. This lag phase determines the end of the herbaceous phase of the berry. The seeds stock up reserves, expand due to water accumulation and reach their maximal size a few days before *véraison*, the onset of ripening.

Phase III: Maturation phase
Growth resumes during the maturation phase (Phase III) which begins at *véraison*, a milestone from which the berry softens quickly and the skin becomes coloured in red cultivars or translucent in white cultivars. The berries enlarge, become softer, sweeter and less acidic. The volume of the berry is doubled during this phase, which lasts around 45 days. Mesocarp cells expand under vacuole pressure which represents 99% of their volume (Diakou & Carde, 2001). In seeds, reserves accumulation slows down and finally stops at maturity (Egli *et al*., 1998, Cadot *et al*., 2006).

It has been shown that the xylem, which provides water and nutrients, is impeded at the beginning of phase III (reviewed by Ollat *et al*., 2002). As a result, the phloem sap, transported from leaves (source organs), must be responsible for water and sugars import in the berry during phase III. Accumulated in leaves during the first growth phases, potassium (K^+) must be forwarded towards the berries *via* the phloem sap (Blouin & Cruège, 2003).
Interacting with tartaric acid, potassium influences the global pH of the berry (Gawel *et al*., 2000), and indirectly, the taste, the colour and the colour stability of wine.
Anthocyanins are synthesized and accumulated during this phase, conferring colour to the skin of red cultivars.
PAs oxidation triggers a seed coat colour shift from light green to dark brown (Kennedy *et al*., 2000, Cadot *et al*., 2006).
The majority of compounds accumulated during phase I are found in lower concentrations (expressed *per* berry mass) during phase III due to the berry mass increase which triggers their dilution (e.g.

tartaric acid, Conde *et al*., 2007). In contrary, malic acid is not simply diluted. An increase of tonoplast leakage was observed during ripening (Terrier *et al*., 2001), resulting in malate decompartmentation. This organic acid may be further metabolized into sugars. It could also be used for mitochondrial respiration (Koch & Allweldt, 1978). The decrease in the amount of malic acid causes an increased vacuolar pH (up to 3.5 at maturity, Ruffner, 1982).
The grape berry is a sink organ able to accumulate sugars routed from photosynthetic source organs (e.g. leaves, stems) *via* the phloem (Kingston-Smith, 2001, Boss & Davies, 2001). Hexoses (fructose, glucose) accumulate in the vacuole of the mesocarp cells from *véraison*. Cell wall and vacuolar invertases are enzymes responsible for conversion of sucrose, the main translocated sugar, into fructose and glucose (Davies & Robinson, 1996, Zhang *et al*., 2006, Boss & Davies, 2009). Several sugar transporters called Hexose transporter (*VvHT*) must be involved in their transport through the tonoplast (reviewed by Conde *et al*., 2007, Agasse *et al*., 2009).
Aroma precursors and compounds (terpenes, norisoprenoids, esters, and thiols) are mainly synthesized at the end of phase III (Lund & Bohlmann, 2006).

The end of the ripening phase was estimated at 120 days after flowering (Carbonneau *et al*., 2007). The berries can further wither due to the concomitant obstruction of the phloem and stomatal water loss (McCarthy & Coombe, 1999).

1.3. Hormonal control

Phytohormones produced by seeds: auxin, cytokinins, gibberellins are involved in the control of the development and the cell growth (Symons *et al*., 2006, Conde *et al*., 2007).
The control of the grape berry development results from the complex interaction of several hormonal signalling pathways (Fortes *et al*., 2015). This review highlights 2 groups of phytohormones based on their effect on the maturation: promoter (ethylene, abscisic acid and brassinosteroids) or inhibitor (auxins, cytokinins, gibberellins).

Grape being a non-climacteric fruit, the onset of maturation is not initiated by the typical ethylene peak and respiratory burst found in climacteric fruits. Nonetheless, ethylene is involved in some maturation processes such as accumulation of anthocyanins (El-Kereamy *et al*., 2003, Chervin *et al.*, 2004), expression of alcohol dehydrogenase (Tesnière *et al*., 2006) and reduction of acidity (Chervin *et al*., 2004).
Abscisic acid (ABA) and brassinosteroids, which accumulate post-ripening, are likely involved in the control of the maturation. An exogenous brassinosteroids treatment accelerates the ripening berries (Symons *et al*., 2006). An ABA peak has been reported at *véraison* in red berries (Deluc *et al*., 2009, Gambetta *et al*., 2010). ABA promotes maturation (Wheeler *et al*., 2009) and seed dormancy (Coombe & Hale, 1973), as well as accumulation of hexoses (Davies *et al*., 1997). In the skin, ABA decreases the expression of the biosynthetic genes of the flavan-3-ols at the beginning of the berry ripening (Lacampagne *et al*., 2010) but induces anthocyanin-specific genes (Jeong *et al*., 2004). The simultaneous presence of sucrose and ABA induced anthocyanin accumulation in detached berries cultured *in vitro* (Gambetta *et al*., 2010).

The auxin indole-3-acetic acid (IAA) is known to delay berry size, sugar and anthocyanin content (Davies *et al*., 1997, Böttcher *et al*., 2010, Ziliotto *et al*., 2012). IAA content through berry development is controversial. It has been reported as low during all the berry development (Davies &

Böttcher, 2009, Symons *et al*., 2006) whereas another study showed an IAA peak during the green stage (Zhang *et al*., 2003).

High cytokinins content has been reported in the flesh of immature berries but their concentration decreases at *véraison* and stays low during maturation (Zhang *et al*., 2003). This hormone would favour cell division during the first days of development.

In immature berries, the content of active gibberellins was high, but decreases along the rest of berry development (Zhang *et al*., 2003, Symons *et al*., 2006). Few evidences support that gibberellins could directly control the grape ripening.
The involvement of other phytohormones such as strigolactones is still to be determined.

2. Phenolic compounds

Phenolic compounds are ubiquitous molecules widely encountered in vascular plants and include molecules with a phenolic ring, i.e. a benzyl ring with at least one hydroxyl group. Despite their diversity, these molecules have a common origin, the aromatic amino acid phenylalanine. Its deamination allows the biosynthesis of cinnamic acid, the entry point into the phenylpropanoid pathway. This pathway is generator of numerous phenolic compounds. These secondary metabolites are crucial in defence mechanisms against biotic and abiotic stress, reproduction and dissemination in the plant kingdom.
In planta, most of the phenolic compounds are associated with a sugar (glycosylated form). However, their carbon structure without sugar (aglycone) allows to classify those metabolites into different groups and to distinguish phenolic acids: (i) hydroxybenzoic acid derivatives (C1-C6 structure) or (ii) hydroxycinnamic acid derivatives (C3-C6), stilbenes (C6-C2-C6) and flavonoids (C6-C3-C6) (Fig. 3).
The grapevine has the intrinsic property of being rich in phenolic compounds, including flavonoids of oenological interest. The most abundant are, in descending order, flavan-3-ols (PAs), anthocyanins, and HCAs (Adams, 2006). Flavan-3-ols and anthocyanins are flavonoids whereas HCAs are cinnamic acid derivatives. The "phenolic maturity" is considered as a maturity level that takes into account both content and extractability of phenolic compounds (Glories, 1998). However, this concept lacks concrete criteria and is rather controversial (Fournand & Moutounet, 2008).
There is an important variability concerning phenolic compounds concentration in the grape berry between cultivars. Their structure of free form and derivatives (glucosylated, methylated, acylated, polymerized forms), spatiotemporal accumulation pattern in berry and the parameters that could influence them are described in this part. We describe those phenolic compounds in plants and then focus on grapevine. Other major classes of phenolic compounds such as coumarins, hydrolysable tannins, monolignols and lignin will not be detailed. Hydrolysable tannins, built from GA, are absent in grape berry and their presence in wine results of barrel aging (Puech *et al*., 1999). Otherwise, lignin formed in the grape berry seed coat is not extracted during fermentation (Keller, 2015).
The bibliographic study focuses on key hydroxybenzoic acids (HBAs), hydroxycinnamic acids (HCAs), stilbenes and flavonoids found in grape berry, with special attention paid to the GA and flavan-3-ols.

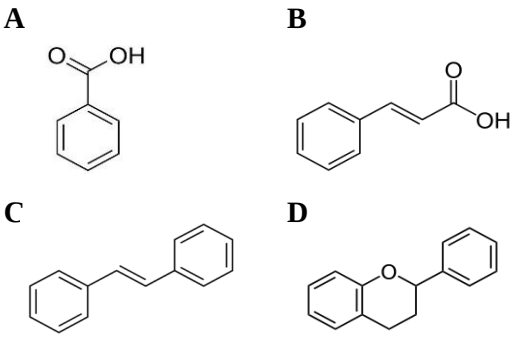
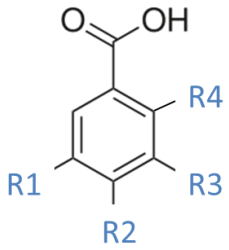


Figure 3. Carbon structure of phenolic compounds found in grape berry.
A. C1-C6 structure of benzoic acid derivatives. **B**. C3-C6 structure of cinnamic acid derivatives. **C**. C6-C2-C6 structure of stilbenes. **D**. C6-C3-C6 structure of flavonoids.



Name	**R1**	**R2**	**R3**	**R4**
Gallic	OH	OH	OH	H
Protocatechuic	H	OH	OH	H
Gentisic	OH	H	H	OH
p-hydrobenzoic	H	OH	H	H
Vanillic	H	OH	OCH_3	H
Syringic	OCH_3	OH	OCH_3	H
Salicylic	H	H	H	OH

Figure 4. Structure of the major hydroxybenzoic acids found in grape berry.

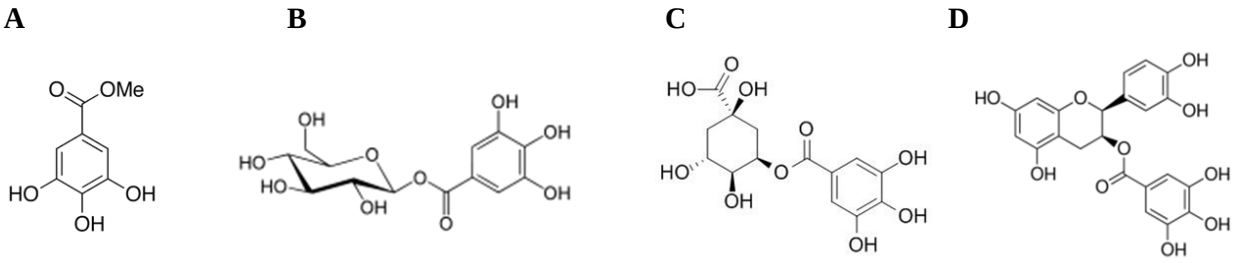


Figure 5. Examples of gallic acid derivatives found in plants.
A. Gallicin. **B**. β-glucogallin. **C**. 5-*O*-galloylquinic acid. **D**. Epicatechin 3-*O*-gallate.

2.1. Phenolic acids

2.1.1. Hydroxybenzoic acids

These molecules are hydroxylated derivatives of benzoic acid and have in common a carbon structure C1-C6. The hydroxylation/methylation pattern of the aromatic ring allows to distinguish these molecules (Fig. 4). In plants, the most common HBAs are usually gallic, protocatechuic, gentisic, *p*-hydroxybenzoic, vanillic, syringic and salicylic acids (Rentzsch *et al*., 2009).

HBA content in grape berry pericarp has been investigated in Cabernet Sauvignon (Chen *et al*., 2006). Gallic, protocatechuic and gentisic are the most accumulated HBAs whereas *p*-hydroxybenzoic, vanillic, syringic acids accumulate at lower contents.

HBAs can be precursors of hormones, cofactors, defense, aroma or flavor compounds (Widhalm & Dudareva, 2015). Salicylic acid is a key molecule for signal perception and transduction, notably involved in plants systemic acquired resistance (An & Mou, 2011). Salicylic acid (free form and glucosylated) has also been studied for its involvement in thermotolerance regulation in grapevine leaves (Wang *et al*., 2005).

In berry pericarp, a first accumulation of HBA has been observed in the early stage of berry development and the amount is kept constant between 40 and 60 days after flowering (daf, Chen *et al*., 2006). Then, it rises again from 60 to 80 daf and tends to decrease after.

Leaf removal before blooming triggered a 50% reduction of total HBAs in Istrian Malvasia grape juice compared to the control condition (Bubola *et al*., 2012).

2.1.1.1. Gallic acid and derivatives

GA is a trihydroxybenzoic acid, also called 3,4,5-trihydroxybenzoic acid. This molecule is widely used as a standard in quantitation of phenolic compounds, the concentration being then expressed as GA equivalents. GA is found in plant derived foods such as tea, honey (Pyrzynska & Biesaga, 2009), pomegranate products (Qu *et al*., 2012), vinegar (Cerezo *et al*., 2008).

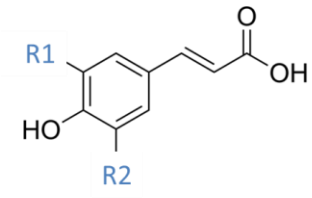
GA is mainly known as a precursor for the biosynthesis of hydrolysable tannins as demonstrated in *Quercus* sp. and *Rhus typhina* (Niemetz & Gross, 2005). It can also be esterified by flavan-3-ols in some plant species like grape (*Vitis vinifera*), tea plant (*Tea sinensis*), and persimmon (*Diospyros kaki*). When not associated with tannins, GA can be found in free, methylated (GA methyl ester also called gallicin) and glucosylated forms (GA glucose ester called β-glucogallin, β-G), or can be esterified by quinic acid to form galloylquinic acids found in some species: tea plant (Jiang *et al*., 2013), strawberry (Vallarino *et al*., 2015) (Fig. 5, Supplemental Table 1). Rarer GA derivatives identified in plants have been listed in Supplemental Table 2. Few studies have simultaneously reported GA content from different plant species. Due to the diversity of the methods used for GA extraction and quantification, and the organs/tissues examined, data comparison between different studies is difficult.

Table 1. GA content in grape organs.

Species	**Cultivar**	**Tissue**	**Content (**nmol.g^{-1})	**Reference**
Vitis rotundifolia		Seeds	194-676	Pastrana-Bonilla *et al*., 2003*
		Leaves	359-1099	
Vitis rotundifolia		Seeds	5819	Yilmaz &Toledo, 2004 **
Vitis vinifera	Chardonnay	Seeds	882	
Vitis vinifera	Merlot	Seeds	588	
Vitis vinifera	Chardonnay	Skin	294	
Vitis vinifera	Merlot	Skin	176	

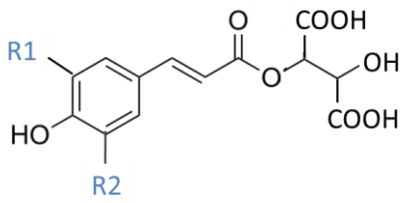
Original data provided in mg/100g have been converted in nmol.g^{-1} using M=170.12 g.mol^{-1} for gallic acid molecular weight.
*: Data expressed by gram of fresh weight.
**: Data expressed by gram of dry weight.

A



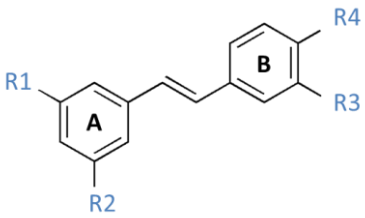
Name	**R1**	**R2**
Caffeic	OH	H
Ferulic	OCH_3	H
p-coumaric	H	H
Sinapic	OCH_3	OCH_3

B



Name	**R1**	**R2**
Caftaric	OH	H
Fertaric	OCH_3	H
Coutaric	H	H

Figure 6. Structure of major hydroxycinnamic acids of grape berry.
A. Free form of hydroxycinnamic acids. **B**. Hydroxycinnamoyl tartrates.



Name	**R1**	**R2**	**R3**	**R4**
trans-resveratrol	OH	OH	H	OH
trans-piceid	OH	O-Glc	H	OH
trans-pterostilbene	OCH_3	OCH_3	H	OH
trans-piceatannol	OH	OH	OH	OH

Figure X. Grapevine simple stilbenes.
Glc: glucose.

GA has been identified as a root-secreted allelochemical in *Phragmites australis*, phytotoxic for *Arabidopsis thaliana* and the other tested species (Rudrappa *et al*., 2007).

In grapevine, GA is mainly accumulated as galloyl moieties acylating flavan-3-ols and few studies have evaluated free or glucosylated GA content. In *Vitis vinifera* and *rotundifolia* (muscadine grape), GA was found in the seeds and leaves predominantly (Table 1).
In grape pericarp, GA is the most abundant HBA (Chen *et al*., 2006). Two peaks of accumulation have been observed along berry development in pericarp: the first one at 30 daf (~100 µg per berry) and the second one at 80 daf (~140 µg per berry).

2.1.2. Hydroxycinnamic acids

Cinnamic acid is a key intermediate between the shikimate pathway and phenylpropanoids (cf. Biosynthetic pathways). Cinnamic acid comprises an aromatic ring (C6) on to which is grafted a chain of 3 carbons (C3) terminated by a carboxylic group (Fig. 6, A). In plants, the most common cinnamic acid derivatives are caffeic, ferulic, *p*-coumaric and sinapic acids, the last being absent or accumulated at very low levels in grape. Their content in different fruits, vegetables and beverages frequently consumed was recently reviewed (El-Seedi *et al*., 2012).
Different HCA derivatives are found in plant kingdom: sinapoyl malate and sinapoylcholine in Arabidopsis, chlorogenic acids (e.g. in Solanaceae: tobacco (Ncube *et al*., 2014) and tomato (Niggeweg *et al*., 2004)), phaselic acid in red clover (Sullivan & Zeller, 2012).
In grape berry, HCAs are accumulated as hydroxycinnamoyl tartrates and anthocyanins, the most abundant phenolic compounds after flavan-3-ols and anthocyanins. Tartaric acid acylation with caffeic, ferulic or *p*-coumaric acids yields caftaric, fertaric and coutaric acids, respectively (Ribéreau-Gayon, 1964, Singleton *et al*., 1978) (Fig. 6, B).

HCA derivatives could be involved in defence against different (a)biotic stresses (Thipyapong *et al*. 2004, Lallemand *et al*., 2012).

In grape berry, hydroxycinnamoyl tartrates accumulate in the cells of the hypodermis in the skin and in placental cells of pulp. They are the major phenolic compounds found in berry pulp (Braidot *et al*., 2008). Their biosynthesis occurs during phase I of the berry development (Romeyer *et al*., 1983).
HCAs can also acylate anthocyanins on their glucosyl part.

The skin concentration of tartaric acid and anthocyanins acylated with hydroxycinnamoyl moieties varies depends on the cultivar (Boursiquot *et al*., 1986, Roggero *et al*., 1986, Ferrandino *et al*., 2012).

2.2. Stilbenes

Stilbenes are non-flavonoids phenolic compounds having a C6-C2-C6 carbon structure (1,2-diphenylethylene, Fig. 7). Some vascular plant families are known to accumulate high stilbene content, especially Fabaceae (e.g. peanut), Pinaceae (e.g. pine), Vitaceae and Poaceae (e.g. sorghum).

Stilbenes found in *Vitis vinifera* are formed from a common molecule, *trans*-resveratrol (3,5,4'-trihydroxy-*trans*-stilbene). Resveratrol derivatives are also found in the berry: the piceatannols

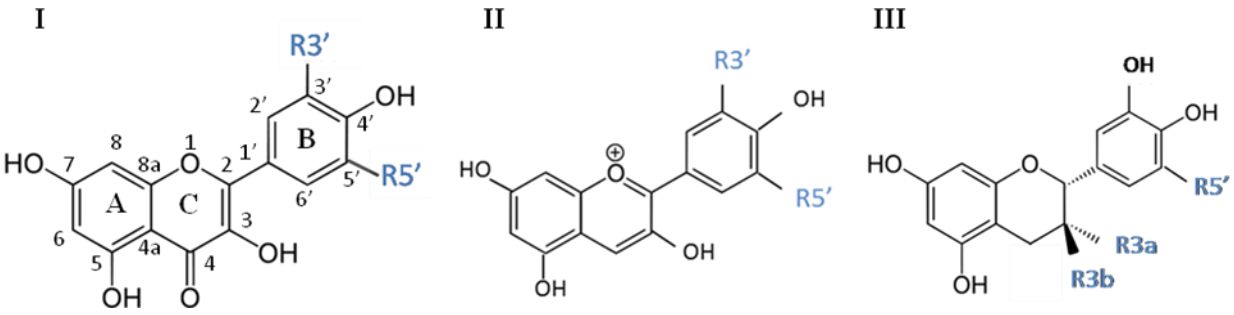


Figure 8. Common structure of flavonoids.
I. Flavonol. **II**. Anthocyanidin. **III**. Flavan-3-ol.

(hydroxylated forms, Bavaresco *et al*., 2002) and piceids (glycosylated forms, Waterhouse & Lamuela-Raventos, 1994), pterostilbens (methylated forms). Resveratrol polymerization is catalyzed by a peroxidase capable of forming viniferins (Barceló *et al*., 2003). Thus, resveratrol dimers (δ- or ε-viniferins), trimers (α-viniferin), tetramer (β-viniferin) or more polymerized form (γ-viniferin) have been identified (Pezet *et al*., 2003, Vitrac *et al*., 2005, Jean-Denis *et al*., 2006). Eighteen different stilbenes have been identified in grape (Flamini *et al*., 2013). The main stilbenes accumulated in grapevine are resveratrol, piceid, piceatanol and resveratrol dimers (Bavaresco *et al*., 2002, Vitrac *et al*., 2005).

These molecules have been particularly studied for their role in defence against major grapevine pathogens (*Plasmopora viticola*, *Botrytis cinerea*) and protection against light exposure. Otherwise, resveratrol could be involved in the beneficial effects of wine on health (Baur & Sinclair, 2006).

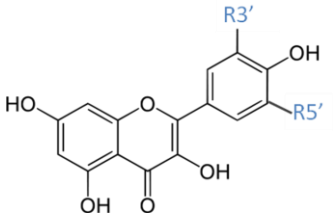
Stilbenes are mainly found after *véraison* in skin, and to a lesser extent in pulp (Gatto *et al*., 2008).

Their content is variable in *V.vinifera* cultivars (Gatto *et al*., 2008). It has been shown that stilbenes accumulate in grape berry under UV light (Petit *et al*., 2009). Exposure to light promotes the isomerisation of *trans*-resveratrol to *cis*-resveratrol (Roggero & Garcia-Parrilla, 1995).
Their accumulation is induced by mildew (Adrian *et al*., 1997, Pezet *et al*., 2004, Fung *et al*., 2008) and *Botrytis cinerea* (Bavaresco *et al*., 2007) infection. Their localization in berry skin coincides with their role in protection against pathogen attack (Fornara *et al*., 2008).

2.3. Flavonoids

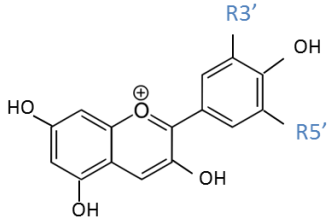
More than 6,000 molecules belonging to the family of flavonoids have been identified to date in Pteridophyta, Bryophyta, Gymnosperms and mostly in Angiosperms (Harbone & Williams, 2000). These molecules have a very high structural diversity and are involved in numerous biological roles. Flavonoids can be defined as "hybrid compounds" as their carbon structure C6-C3-C6 results from the combination of the phenylpropanoid pathway C6-C3 carbon structure, and the route of polyketides which provides a second C6 cycle (Quideau *et al*., 2011). Indeed, flavonoid biosynthesis begins with the association of the 4-coumaroyl-CoA from the phenylpropanoid pathway and malonyl-CoA from the fatty acids pathway. The flavonoids consist of 2 phenolic rings (A and B) linked by a central heterocycle (C) (Fig. 8). The oxidation degree of cycle C allows to distinguish the three main flavonoid classes found in grape berry: flavonols, anthocyanins and flavan-3-ols. These have in common a hydroxyl (–OH) group at positions 3, 5, 7 and 4'. Additional substitution by hydroxyl and/or methyl (-CH_3) groups on the C3’ and C5’ at B ring of the carbon structure distinguishes the different aglycones.

Many biological roles are attributed to flavonoids found in vascular plants. They are involved in important physiological functions in different organs of the plant, and can accumulate in response to (a)biotic stresses.



Name	R3'	R5'
Kaempferol	H	H
Quercetin	OH	H
Isorhamnetin	OCH_3	H
Myricetin	OH	OH
Laricitrin	OCH_3	OH
Syringetin	OCH_3	OCH_3

Figure 9. Structure of the major flavonols found in grape berry.



Name	R3'	R5'
Cyanidin	OH	H
Peonidin	OCH_3	H
Delphinidin	OH	OH
Malvidin	OCH_3	OCH_3
Petunidin	OH	O CH_3

Figure 10. Major anthocyanin aglycones found in grape berry.

2.3.1. Flavonols

Flavonols exhibit a carbonyl function at C4 and a double bond between C2 and C3 (Fig. 8, I). In plants, they are mainly accumulated as glycosides. The nature of the sugar bound to carbon 3 extends their diversity.
The main flavonol found in grape berry is quercetin, followed by myricetin, kaempferol, isorhamnetin and laricitrin (Mattivi *et al*., 2006, Fig. 9). The flavonols are present in the berry as glucoside, galactoside, glucuronide and diglycosides. Quercetin 3-*O*-glucoside and 3-*O*-glucuronide are the most abundant flavonols in the grape berry (Cheynier & Rigaud, 1986; Price *et al*., 1995).

Flavonols confer protection against UV radiation (Smith & Markham, 1998). They are involved in the yellow colouring of certain organs such as flowers and help attract pollinators, and contribute to the pollen fertility (Taylor & Jorgensen, 1992). Those signal molecules are also involved in the interaction of the plant with microorganisms (Koes *et al*., 1994). Rutin exhibits insecticidal properties (Simmonds, 2003).When associated with anthocyanins, they promote copigmentation and thus stabilize the colour of anthocyanins (Osawa, 1982). Flavonols negatively regulate the transport of a phytohormone, auxin (Peer & Murphy, 2007) and the development of the pollen tube (Thompson *et al*., 2010).

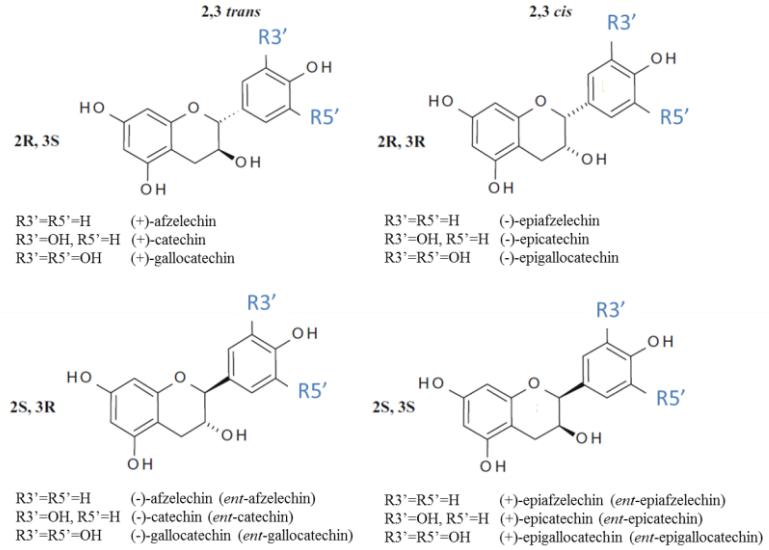
Flavonols have been found in grape berries, leaves, tendrils, buds, inflorescences, anthers (Downey *et al*., 2003a) and stems (Souquet *et al*., 2000). In grape berry, they accumulate in skin (Souquet *et al*., 1996), and in pulp to a lesser extent (Pereira *et al*., 2006). More particularly, flavonols are accumulated in the upper epidermis of the berry. Therefore, the varieties having the higher skin/berry volume ratio provide higher concentrations of flavonols (McDonald *et al*., 1998). They are accumulated during two development stages: at flowering and 3-4 weeks after *véraison* (Downey *et al*., 2003a).

Grape berries exposed to sunlight have a higher concentration of flavonols than those shaded, in accordance with their role as UV shield (Price *et al*., 1995). Flavonol pattern differs among grapevine cultivars and can be used for taxonomical classification (Mattivi *et al*., 2006). The comparison of mature berries from 91 cultivars reveals that myricetin and its derivatives (laricitrin, syringetin) are red cultivars specific. Quercetin is the major flavonol in white culivars but is also abundant in red cultivars.

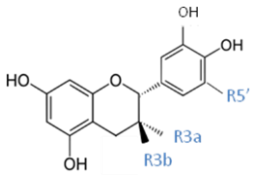
2.3.2. Anthocyanins

Anthocyanins are composed of an aglycone part (anthocyanidin) bound to a sugar. The aglycone is usually presented as the flavylium cation (2-phenylbenzopyrilium, Fig. 8, II) but this form is predominant in very acidic conditions only. In vascular plants, a high variability exists regarding the sugar moieties (e.g. galactose, rhamnose, arabinose, Andersen & Jordheim, 2006) mostly linked as *O*-glycoside and their position on the carbon structure (mainly on carbon 3). *C*-glucoside bonds were also identified for anthocyanins (Saito *et al*., 2003). Moreover, anthocyanin diversity is extended by acyl moieties linked to the sugar.

A



B



Name	Abb.	R3a	R3b	R5'	C2-C3 configuration
(+)-catechin	-	H	OH	H	2R, 3S (*trans*)
(-)-epicatechin	EC	OH	H	H	2R, 3R (*cis*)
(-)-epicatechin gallate	ECG	GA	H	H	2R, 3R (*cis*)
(+)-gallocatechin	GC	H	OH	OH	2R, 3S (*trans*)
(-)-epigallocatechin	EGC	OH	H	OH	2R, 3R (*cis*)

Figure 11. Structure of flavan-3-ol monomers.
A. Non-galloylated monomers. **B**. Major monomers found in grapevine.

In *Vitis vinifera*, the 5 major aglycones are cyanidin, delphinidin, malvidin, peonidin, and petunidin (Fig. 10). The groups found on the aglycone influence its absorption capacity and thus modulate its colour from orange (cyanidin) to purple (malvidin).
In anthocyanins of *Vitis vinifera*, the glycoside moiety is a glucose molecule linked by *O*-glycosylation on carbon 3. Malvidin-3-*O*-glucoside is often the main anthocyanin accumulated in grape berry.The sugar can be acylated with acetic, *p*-coumaric or caffeic acid (e.g. Adams, 2006, Pinelo *et al*., 2006).

The absorption capacity of anthocyanins in the visible light confers colour to different organs of the plant. In the floral organs and fruits, colour attracts pollinators and promotes seed dispersal. Anthocyanins impart colour ranging from red to purple as the pH increases. *In planta*, anthocyanins colour is stabilized by vacuolar pH, association with other molecules (copigmentation), or metals (reviewed by Yoshida *et al*., 2012).

Anthocyanins are found in the subcutaneous tissues of the skin of red grapes and in the pulp of certain cultivars called "teinturiers". Their biosynthesis and accumulation in the berry begins at *véraison*.

Some varieties, such as Pinot Noir, are devoid of acylated anthocyanins (Mazza *et al*., 1999). The anthocyanin concentration varies according to *Vitis vinifera* varieties and *Vitis* subspecies. Anthocyanins content was decreased by shading in Cabernet-Sauvignon berries (Jeong *et al*., 2006) but stable in Syrah berries (Downey *et al*., 2004). High temperatures (>30°C) decreased anthocyanins content, probably due to their degradation in red grape berries (Yamane *et al*., 2006, Mori *et al*., 2005). Moreover, water stress increased anthocyanins content in berries (Castellarin *et al*., 2007, Ollé *et al*., 2011).

2.3.3. Flavan-3-ols

Flavan-3-ols are found as monomers and polymers (condensed tannins or PAs) in many fruits and food products (wine, tea, chocolate) to which they confer astringency (drying sensation in the mouth). Indeed, condensed tannins are able to interact with proteins and precipitate them. Those molecules were used as tanning agents for the production of leather from animal hides. When heated in acidic conditions, PAs yield anthocyanidins, which explains the name "proanthocyanidins". PAs are commonly found in woody species but their presence in herbaceous plants can be restricted to some organs like seed coat (e.g. *Arabidopsis thaliana, Medicago truncatula*).

Concerning the structure of monomers (Fig. 8, III), the asymmetry (*cis* or *trans* configuration) of C2 and C3 carbons accounts for the presence of stereoisomers of flavan-3-ols in plant species. The prefix "epi" means molecules with *cis* configuration of C2 and C3 carbons (2R, 3R or 2S, 3S). Furthermore, the hydroxylation level of the B ring differentiates afzelechins, catechins and gallocatechins. Flavan-3-ols can be acylated with GA at the –OH group of C3 to form galloylated flavan-3-ols (e.g. epicatechin 3-*O*-gallate, ECG). This acylation reaction is called galloylation. The presence of galloyl moieties on flavan-3-ols is restricted to few plant species (Supplemental Table 2). Extreme values, close to 100%, have been reported for bayberry leaves PAs (Yang *et al*., 2011).
(+)-catechin, its isomer (-)-epicatechin (EC), (-)-epigallocatechin (EGC) and to a lesser extent (-)-epicatechin-3-*O*-gallate (ECG) are the main forms of grape flavanol-3-ol monomers (Su & Singleton,

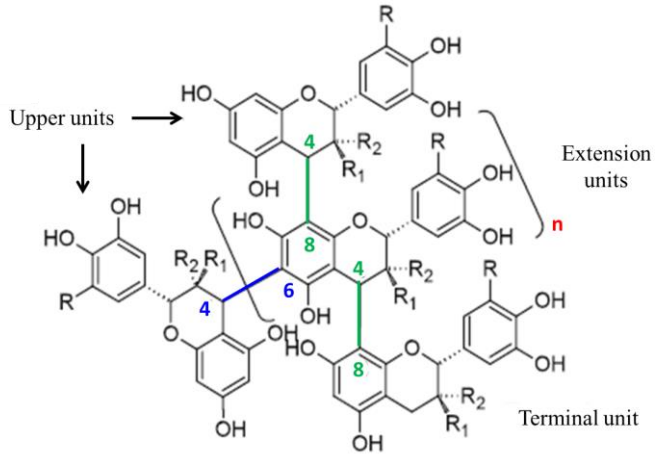


Figure 12. Structure of B-type proanthocyanidins.
n: number of extension units. In green: C4-C8 bond, in blue: C4-C6 bond. This figure is adapted from Huang *et al*., 2012

Table 2. Characteristics of PAs in grapevine (*Vitis vinifera*) tissues and Arabidopsis (*Arabidopsis thaliana*) seeds. Important variations can occur between genotypes (adapted from Huang *et al*., 2012).

Species	**Tissue**	**mDP**	**Major TU**	**Major EU**	**% trihydroxylated B-ring**	**%G**
Grapevine	Skin	20–50	C	E(G)C	+++ (5–40)	+ (1–5)
	Pulp	20–30	C/EC	E(G)C	+ (2–15)	+ (1–5)
	Seed	3–16	C/EC	EC	0	+++ (10–20)
	Leaves	10–50	C/EC	E(G)C	++ (20)	++ (5)
Arabidopsis	Seed	~5	EC	EC	0	0

Abbreviations: **mDP** : mean polymerization degree, **TU** : terminal unit, **EU** : extension unit, **%G**: galloylation percentage, C: (+)-catechin, EC: (-)-epicatechin, E(G)C: (-)-epi(gallo)catechin.
Data from Prieur *et al*., 1994, Bogs *et al*., 2005, Cheynier *et al*., 1998, Monagas *et al*., 2003, Tesnière *et al*., 2006, Souquet *et al*., 2006, Routaboul *et al*., 2006, Mané *et al*., 2007 and Verriès *et al*., 2008.

1969, Figure 11). (+)-gallocatechin (GC) is present in trace amounts. Recently, flavan-3-ols monoglycosides were detected in seeds (Delcambre & Saucier, 2012).

As in the majority of plant species, grape PAs are predominantly of B type, i.e. formed mainly by C4-C8 or C4-C6 bonds between monomers (Figure 12). PAs chains can be depolymerised by heating under acidic conditions. In presence of a nucleophile (e.g. phloroglucinol), the depolymerised subunits form adducts. Then, HPLC analysis allows the determination of different PAs characteristics and the proportion of their constitutive units. The mean degree of polymerization (mDP) is defined as the mean number of subunits constituting the PAs chains. The main characteristics of PAs in the 2 most studied plants for their biosynthesis (grapevine and Arabidopsis) are compared in the Table 2. Arabidopsis PAs are only found in seeds and have a lower mDP (~5). This plant is devoid of trihydroxylated and galloylated flavan-3-ols. In grape berry, mDP usually ranges between 2 and 50, according to the tissue. PAs have a higher mDP but a lower %G in the skin than in seeds. %G varies between 10 and 20% in the seeds and between 1 and 5% in the skin according to the cultivar. PA subunits with trihydroxylated B-ring are absent in seeds.

PAs are colourless flavonoids that turn brown under the effect of oxidation. Their accumulation in the seed coat influences the germination of *Arabidopsis thaliana* (Debeaujon *et al*., 2000, 2001). Accumulation of PAs could protect organs where they accumulate against water stress (Hernández *et al*., 2006) and, notably at high altitudes, against UV-B radiation (Alonso-Amelot *et al*., 2007). They are also involved in the defence against various pathogens and herbivores (Peters & Constabel, 2002, Dixon *et al*., 2005, Miranda *et al*., 2007). An increased galloylation rate and *cis* configuration of PAs was observed in tea leaves after infection by *Exobasidium vexans* (Nimal Punyasiri *et al*., 2004).

PAs are the most abundant phenolic compounds in the grape berry. They are also present in other parts of the plant: wood (Boukharta *et al*., 1988), leaves (Bogs *et al*., 2005) and stems (Souquet *et al*., 2000). In skin, they accumulate in hypodermic layers. In seeds, they are found in epiderm, outer and inner teguments (internal layers) and their biosynthesis occurs during early development (Cadot *et al*., 2006). The pulp is poor in PAs compared to skin and seeds (Mané *et al*., 2007, Verriès *et al*., 2008).
Flavan-3-ols are stored in the vacuole and in the plant cell wall (Matthews *et al*., 1997, Kennedy *et al*., 2001). In skin, flavan-3-ols are synthesized during the early phases of development of the berry (green stage) with a maximum content reached before *véraison* (Kennedy *et al*., 2001, Downey *et al*., 2003b, Verriès *et al*., 2008). Their content remains stable after *véraison* in skin when expressed on a *per* berry basis (Fournand *et al*., 2006). In seeds, their concentration is maximal 2 weeks after *véraison* (Downey *et al*., 2003b, Bogs *et al*., 2005).

Contrary to flavonols and anthocyanins, abiotic stresses seem to have little impact on PAs content in grape berry. PA composition is few impacted by water stress (Ojeda *et al*., 2002, Kennedy *et al*., 2002, Ollé *et al*., 2011), and light exposure slightly increases mDP and % trihydroxylated flavan-3-ols in skin (Downey *et al*., 2004, Cortell & Kennedy, 2006).

The concentration and structure of PAs can vary within the same tissue according to the cultivar. It has been shown that the skin PA concentration is higher in Monastrell berries than in Cabernet-Sauvignon berries and Syrah cultivars harvested at a similar stage of development (Busse-Valverde *et al*., 2010). However, Syrah berry skin exhibits the highest %G (6.6%), about 4 times higher than in the skin of Monastrell and Cabernet Sauvignon berries.

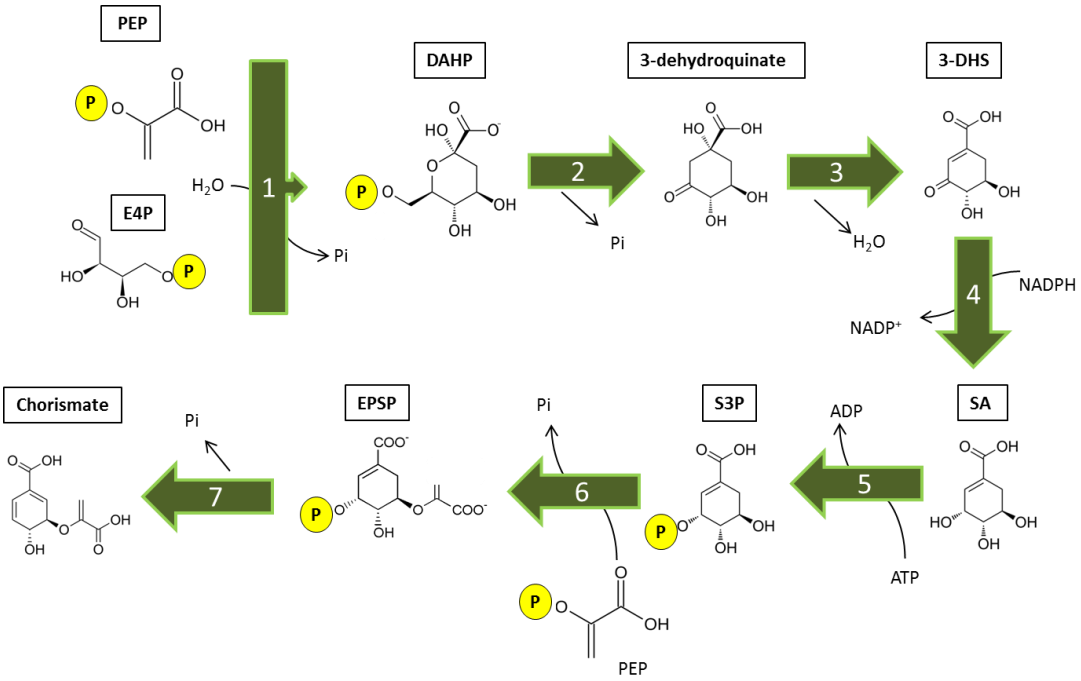


Figure 13. Shikimate pathway in plants.
Metabolites: PEP: phophoenolpyruvate, E4P: D-erythrose 4-phosphate, DAHP: 3-deoxy-D-arabino-heptulosonate 7-phosphate, 3-DHS: 3-dehydroshikimate, SA: shikimic acid, S3P: shikimate 3-phosphate, EPSP: 5-*enol*pyruvylshikimate-3-phosphate. P: Phosphate.
Enzymes: **1**: 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase, **2**: 3-dehydroquinate synthase, **3**: 3-dehydroquinate dehydratase, **4**: shikimate dehydrogenase, **5**: shikimate kinase, **6**: 5-enolpyruvylshikimate 3-phosphate synthase, **7**: chorismate synthase.

Table 3. Enzymes involved in the shikimate pathway in plants and number of isogenes in grapevine.

Enzyme	**Abbreviation**	**EC name**	**Grapevine isogenes number**
3-deoxy-D-arabino-heptulosonate 7-phosphate synthase	DAHPS	EC 2.5.1.54	3
3-dehydroquinate synthase	DHQS	EC 4.2.3.4	1
3-dehydroquinate dehydratase/shikimate dehydrogenase	DQD/SDH	EC 4.2.1.10/EC 1.1.1.25	4 (2)
Shikimate kinase	SK	EC 2.7.1.71	2
	SKL1		1
	SKL2		1
5-*enol*pyruvylshikimate 3-phosphate synthase	EPSPS	EC 2.5.1.19	2 (2)
Chorismate synthase	CS	EC 4.2.3.5	2

The number of grapevine isogenes was published by Tohge *et al*., (2013). The number of isogenes between brackets indicates tandem gene duplication event identified in the same study.

3. Biosynthetic pathways

The metabolic pathways described in this section, i.e. the shikimate pathway, biosynthesis pathway of phenylalanine, the phenylpropanoid pathway and flavonoid pathway are necessary for the biosynthesis of HBAs, HCAs and flavonoids.

3.1. Shikimate pathway

The shikimate pathway takes its name from one of its intermediate, shikimate or shikimic acid (SA). This compound was isolated from the Japanese star anise (*Illicium anisatum*), a tree called Shikimi-no-ki in Japanese (Eijkman, 1885). Shikimate is used for the chemical synthesis of oseltamivir, an antiviral better known as Tamiflu®.

The shikimate pathway is common for microorganisms (bacteria, fungi and some protists) and vascular plants, and links the biosynthesis of carbohydrates to the aromatic amino acids (reviewed by Tzin & Galili, 2010). Being deprived of the shikimate pathway, animals do not have the ability to produce aromatic amino acids which are essential to their metabolism. The absence of this pathway in animals has favoured the use of glyphosate, an herbicide that targets an enzyme of the shikimate pathway (Duke & Powles, 2008).

The shikimate pathway leads to the production of chorismate, the precursor of vitamin K1 (phylloquinone), B9 (folic acid), and salicylic acid. Chorismate is also the precursor of the biosynthesis of 3 aromatic amino acids: L-tryptophan, L-tyrosine and L-phenylalanine. When not used for protein biosynthesis, these compounds are precursors for molecules involved in various biological processes in plants (reviewed by Maeda & Dudareva, 2012).
L-tryptophan is a precursor for biosynthesis of alkaloids, phytoalexins, glucosinolates, and auxin. L-tyrosine allows the biosynthesis of other classes of alkaloids, betalains and quinones. L-phenylalanine, converted by the phenylpropanoid pathway, is responsible for the biosynthesis of all the phenolic secondary metabolites found in these organisms.
Under normal growth conditions, it was estimated that 20% of the carbon fixed by vascular plants flows through the shikimate pathway (Haslam, 1993). The study of model plants such as *Arabidopsis thaliana* allowed the characterization of all the enzymes in the shikimate pathway. Seven metabolic steps constitute the plant shikimate pathway (Table 3, Fig. 13). However, the shikimate pathway can be subject to branchpoints leading to the formation of numerous benzoic acids, and possibly GA, as recently reviewed by Widhalm & Dudareva (2015).

In the plant kingdom, the shikimate pathway is recognized as plastidic. This hypothesis is supported by the presence of chloroplast targeting signal peptide within the coding sequences of all the enzymes of the pathway (Schmid & Amrhein, 1999). The identification of a cytosolic isoform of 3-dehydroquinate dehydratase/shikimate dehydrogenase in tobacco (*Nicotiana tabacum*, Ding *et al*., 2007) and the detection of isoforms of other enzymes devoid of targeting peptide (Ganson *et al*., 1986,

Table 4. Kinetic parameters of studied shikimate dehydrogenases in plants.

Species	Tissue	Conditions	Km	Reference
Spinacia oleracea (spinach)	Leaves	glycine NaOH (pH 9.5), 20°C	SA: 200 µM $NADP^+$: 8-26 µM	Fielder & Schultz, 1985
Pisum sativum (pea)	Leaves		SA: 1.5 mM $NADP^+$: 32 µM 3-DHS: 210 µM NADPH: 11 µM	Mousdale *et al*., 1987
Capsicum annuum (sweet pepper)	Seedlings	Tris-HCl (pH 9), 25°C	SA: 87 µM $NADP^+$: 17 µM	Diaz & Merino, 1997
Arabidopsis thaliana		pH 8.8, 22°C	SA: 685 µM $NADP^+$: 131 µM	Singh & Christendat, 2006 *
Nicotiana tabacum (tobacco)		glycine–NaOH buffer (pH 9.0)	SA: 130 µM $NADP^+$: 31 µM	Ding *et al*., 2007 *
Juglans regia (walnut)		Tris-HCl (pH 9)	SA: 860 µM $NADP^+$: 860 µM	Muir *et al*., 2011 *
Populus trichocarpa (poplar)		Trizma base-HCl (pH 8.5)	SA : 223.1-345.8 µM	Guo *et al*., 2014 *

*: heterolougous expression.

Mousdale & Coggins, 1985, D'Amato, 1984) suggests that a parallel shikimate pathway could exist in the cytosol.
A phylogenetic study on the shikimate pathway genes revealed the number of isogenes present in sequenced vascular plants genomes, and gene duplication events (Tohge *et al*., 2013).
The characteristics of the enzymes involved in the shikimate pathway in plants are described in this paragraph.

The first step of the shikimate pathway is catalyzed by 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHPS, Fig. 13, step 1). This enzyme catalyzes the aldol condensation of phophoenolpyruvate (PEP) and D-erythrose 4-phosphate (E4P) to form 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP, Suzich *et al*., 1985). The ion Mn^{2+} and thioredoxin under reduced form are essential for its activity (Entus *et al*., 2002). One to 8 isoforms are present in higher plant genomes (Tohge *et al*., 2013).

3-dehydroquinate synthase (DHQS) catalyzes the 2nd step of the shikimate pathway. This monofunctional enzyme converts DAHP into 3-dehydroquinate. Generally, one gene was found in the genome of plants, except for soybean (*Glycine max*) genome with 2 *DHQS* (Tohge *et al*., 2013).

The 3rd and 4th steps of the pathway are catalyzed by the bifunctional enzyme 3-dehydroquinate dehydratase/shikimate dehydrogenase (DQD/SDH). One to seven genes encoding DQD/SDH have been identified in plants. The DQD activity allows the dehydration of 3-dehydroquinate into 3-dehydroshikimate (3-DHS). The SDH domain catalyzes the reduction of 3-DHS into SA using NADPH as a cofactor. DQD activity is 10 times more important than the SDH activity (Fiedler & Schultz, 1985). This enzyme, whose name varies according to studies, was first purified from the chloroplasts stroma of spinach leaves (*Spinacia oleracea*, Fielder & Schlutz, 1985).
It has been extracted from other plant tissues: pea leaves (*Pisum sativum*, Mousdale *et al*., 1987), sweet pepper seedlings (*Capsicum annuum* L., Diaz & Merino, 1997). In each study, the purified protein had a molecular mass around 60 kDa. The kinetic parameters of those purified enzymes have been studied (Table 4). Zn^{2+} and Cu^{2+} ions were strong inhibitors of its activity (Diaz & Merino, 1997). Protocatechuic acid (PCA) was also found to be a competitive inhibitor.
A partial cDNA of DQD/SDH from tobacco (*Nicotiana tabacum*) was cloned and sequenced (Bonner & Jensen, 1994). For the first time, a cDNA including the N-terminal portion (transit peptide) was cloned from tomato (*Lycopersicum esulentum*, Bischoff *et al*., 2001). In addition, the cloning of two homologous cDNAs of different lengths (one being shorter by 19 amino acids) revealed a phenomenon of transcript splicing.
The 3-dimensional structure of the DQD/SDH was resolved in *Arabidopsis thaliana* (Singh & Christendat, 2006). Key amino acids for substrate and cofactor binding have been highlighted.
Kinetic parameters of DQD/SDHs purified or cloned from several plants are reported in Table 4.
More recently, Muir and co-workers (2011) have shown that a SDH from walnut (*Juglans regia*) was able to produce GA from SA and 3-DHS, using $NADP^{+}$ as cofactor.

Shikimate kinase (SK) is involved in the 5th step of the shikimate pathway. This enzyme catalyzes the ATP-dependent phosphorylation of SA into shikimate 3-phosphate (S3P). Its activity requires the presence of a divalent cation such as Mg^{2+} (Koshiba, 1979). One to 3 isoforms are present in the genome of monocots and dicots. It has been demonstrated that the expression of *SKs* is induced at specific stages of development and biotic stress in *Oryza sativa* (Kasai *et al*., 2005). *SK-Like* genes (*SKL1* and *-2*), involved in chloroplast biogenesis in Arabidopsis, appeared from *SK* duplication (Fucile *et al*., 2008). *SKL* orthologs are present in grapevine genome (Tohge *et al*., 2013).

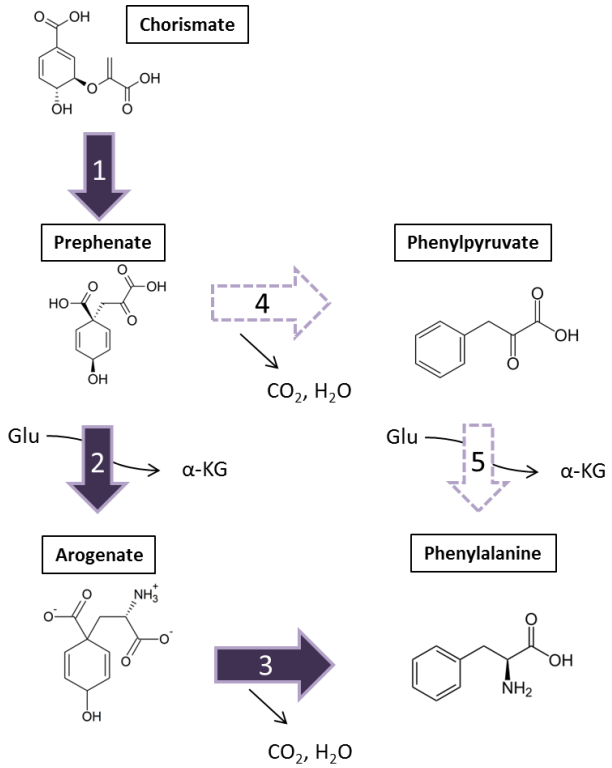


Figure 14. Biosynthesis of phenylalanine in plants.
Glu: Glutamate, α-KG: α-ketoglutarate.
Enzymes: **1**: chorismate mutase (CM), **2**: prephenate aminotransferase (PAT), **3**: arogenate dehydratase (ADT), **4**: prephenate dehydratase (PDT), **5**: phenylpyruvate aminotransferase (PPY-AT).

Table 5. Enzymes involved in the biosynthesis of phenylalanine in plants and number of isogenes in grapevine.
This figure is adapted from Maeda & Dudareva, 2012.

Enzyme	**Abbreviation**	**EC name**	**Produced Amino acid**	**Grapevine isogenes**
Chorismate mutase	CM	5.4.99.5	Tyr/Phe	3
Prephenate aminotransferase	PAT	2.6.1.79	Tyr/Phe	2
Arogenate dehydratase	ADT	4.2.1.91	Phe	2
Prephenate dehydratase	PDT	4.2.1.51	Phe	?
Phenylpyruvate aminotransferase	PPY-AT	2.6.1.5	Phe	?

The number of grapevine isogenes was published by Tohge *et al.*, (2013).

The 6th step of the pathway is catalyzed by 5-*enol*pyruvylshikimate-3-phosphate synthase (EPSPS). The EPSPS allows the condensation of PEP molecule with the S3P to form the 5-*enol*pyruvylshikimate-3-phosphate (EPSP). This enzyme is the target of an herbicide similar to PEP, the glyphosate. Glyphosate treatment decreased the flow of the shikimate pathway (e.g. Amrhein *et al*., 1980, Becerril *et al*., 1989) and increased hydrolysable tannins in birch leaves (Ossipov *et al*., 2003). One to 2 isoforms are present in the genome of plants. In *A. thaliana*, the expression of EPSPS is induced by the attack of *Botrytis cinerea* (Ferrari *et al*., 2007).

Chorismate synthase (CS) catalyzes the 7th and final step of the shikimate pathway and converts EPSP into chorismate. This reaction requires the flavin mononucleotide (FMN), which serves as an electron donor (Macheroux *et al*., 1999). Plant genomes possess from 1 to 3 *CS* isoforms (Tohge *et al*., 2013).

3.2. Biosynthesis of phenylalanine

The biosynthesis pathways of aromatic amino acids is confined in the chloroplast in plants and not found in animals. The different steps of the biosynthesis of aromatic amino acids from chorismate in plants have recently been reviewed (Maeda & Dudareva, 2012). The study focuses on the biosynthesis of phenylalanine, the precursor of the phenylpropanoid/flavonoid pathway.

Chorismate mutase (CM) catalyzes the first of the 2 reactions leading to phenylalanine. CM converts chorismate to prephenate (Fig. 14, step 1). Between 2 and 5*CMs* are present both in monocots and dicots plant genomes (Tohge *et al*., 2013). Prephenate is the common precursor for biosynthesis of phenylalanine and tyrosine.
The prephenate aminotransferase (PAT) converts prephenate to arogenate in presence of glutamate (Fig. 14, step 2). PAT was recently identified in Arabidopsis and petunia (Maeda *et al*., 2011).
Arogenate can be converted into phenylalanine by arogenate dehydratase (ADT, Fig. 14, step 3). ADT was recently cloned and studied in Arabidopsis (Cho *et al*., 2007, Huang *et al*., 2010).

A parallel route for the biosynthesis of phenylalanine from prephenate has also been described in plants. Prephenate dehydratase (PDT) converts prephenate into phenylpyruvate (Fig. 14, step 4). Then, phenylpyruvate aminotransferase (PPY-AT) converts phenylpyruvate into phenylalanine (Fig. 14, step 5). PPY-AT was recently localised in the cytosolic compartment, suggesting the export of chloroplastic phenylpyruvate to the cytosol (Yoo *et al*., 2013).

RNA interference suppression of a petunia *ADT* revealed that phenylalanine is mainly synthesized from arogenate (Maeda *et al*., 2010).

3.3. Phenylpropanoid pathway

3.3.1. Common trunk

The phenylpropanoid pathway is composed of 3 enzymes which allow successive processing of phenylalanine into *p*-coumaroyl-Coenzyme A. This molecule is located at a metabolic crossroad,

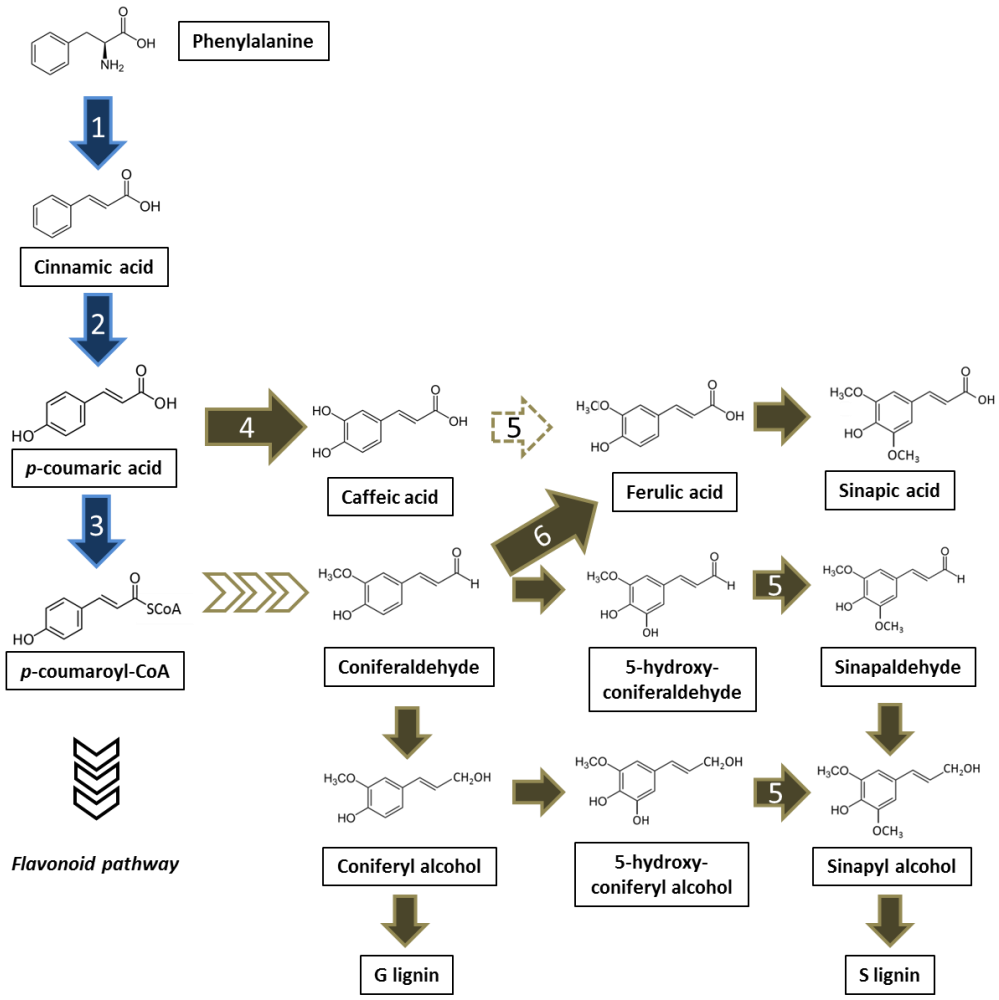


Figure 15. Phenylpropanoid and lignin pathways.
This figure is adapted from Vanholme *et al*., 2010.Phenylpropanoid pathway enzymes (blue): **1**: Phenylalanine ammonia-lyase, **2**: Cinnamate 4-hydroxylase, **3**: 4-coumaroyl:CoA ligase. Lignin pathway enzymes (brown): **4**: *p*-coumarate 3-hydroxylase, **5**: caffeic acid *O*-methyltransferase, **6**: hydroxycinnamaldehyde dehydrogenase. S lignin: syringyl-like lignin structures, G lignin: guaiacyl-like lignin structures.

Table 6. Enzymes involved in the common trunk of phenylpropanoid pathway in plants.

Enzyme	Abbreviation	EC name	Grapevine isogenes number	Reference
Phenylalanine ammonia-lyase	PAL	EC 4.3.1.5	13	Velasco *et al*., 2007
Cinnamate 4-hydroxylase	C4H	EC 1.14.13.11	1	Velasco *et al*., 2007
4-coumaroyl:CoA ligase	4CL	EC 6.2.1.12	1	Velasco *et al*., 2007

being the substrate of numerous enzymes, for the biosynthesis of flavonoids, coumarins, stilbenes, chalcones, lignans and lignins (Boerjan *et al*., 2003).
The phenylpropanoid pathway is organized as a multi-enzyme complex (metabolon) located on the cytosolic side of the endoplasmic reticulum (Hrazdina & Jensen, 1992).

The first enzyme of the phenylpropanoid pathway is phenylalanine ammonia lyase (PAL, Fig. 15, step 1) which is considered as a branch point enzyme between primary and secondary metabolisms (Dixon & Paiva, 1995). This enzyme, which allows the deamination of phenylalanine into cinnamic acid, was discovered from barley seedlings (Koukol & Conn, 1961). Since then, PAL has been widely studied because of its involvement in plant development and response to different stimuli, including in grapevine (Sgarbi *et al*., 2003, Wen *et al*., 2005).

Cinnamate 4-hydroxylase (C4H) converts cinnamic acid into *p*-coumaric acid by hydroxylation (Russell & Conn, 1967, Fig. 15, step 2). C4H is a cytochrome P450-dependent monooxygenase. The single *C4H* gene in *Arabidopsis thaliana* was cloned and characterized (Bell-Lelong *et al*., 1997, Mizutani *et al*., 1997).

4-coumaroyl: CoA ligase (4CL) converts *p*-coumaric acid to 4-coumaroyl-CoA (Heller & Forkmann, 1988, Fig. 15, step 3). The Arabidopsis genome comprises 4 *4CLs* (Costa *et al*., 2005).
PAL, *C4H* and *4CL* have been cloned from *Vitis vinifera* seedlings (Sparvoli *et al*., 1994).
In Cabernet Sauvignon berry, *PAL*, *C4H* and *4CL* exhibit two peaks of activity: the first during the early development (30 daf) and the second after *véraison* (70-80 daf, Chen *et al*., 2006).
In *Vitis vinifera*, *PAL* belongs to a large family of genes whereas a single copy of *C4H* and *4CL* were identified from genomic sequences analysis (Velasco *et al*., 2007). 13 *PAL* isogenes were counted by Velasco and coworkers (2007). 10 isogenes annotated as PAL have been identified from Phytozome database (Wu *et al*., 2014).

5.3.2. Caffeic and ferulic acids biosynthesis

Caffeic acid can be directly formed from *p*-coumaric acid by the *p*-coumarate 3-hydroxylase (C3H, Fig. 15, step 4), which is a cytochrome P450-dependent monooxygenase. C3H has been characterized in Arabidopsis thanks to the *reduced epidermal fluorescence 8* (*ref8*) mutant (Franke *et al*., 2002).

Caffeic acid *O*-methyltransferase (COMT, EC 2.1.1.68) was first propounded as the enzyme responsible for the convertion of caffeic acid into ferulic acid (Fig. 15, enzyme 5). However, this enzyme is rather involved downstream in the phenylpropanoid pathway, converting 5-hydroxyconiferaldehyde and 5-hydroxyconiferyl alcohol into sinapaldehyde and sinapoyl alcohol, respectively (Humphreys *et al*., 1999, Vanholme *et al*., 2010). More recently, the sinapate esters deficient mutant *ref1* of Arabidopsis allowed the identification of an enzyme responsible for ferulic acid biosynthesis. The gene *REF1* encodes a hydroxycinnamaldehyde dehydrogenase (HCALDH) which catalyzes the $NADP^+$-dependent oxidation of coniferaldehyde to ferulic acid (Nair *et al*., 2004, Fig. 15, enzyme 6). This enzyme was also able to produce sinapic acid from sinapaldehyde.
In grapevine, no gene coding for C3H or HCALDH has been characterized.

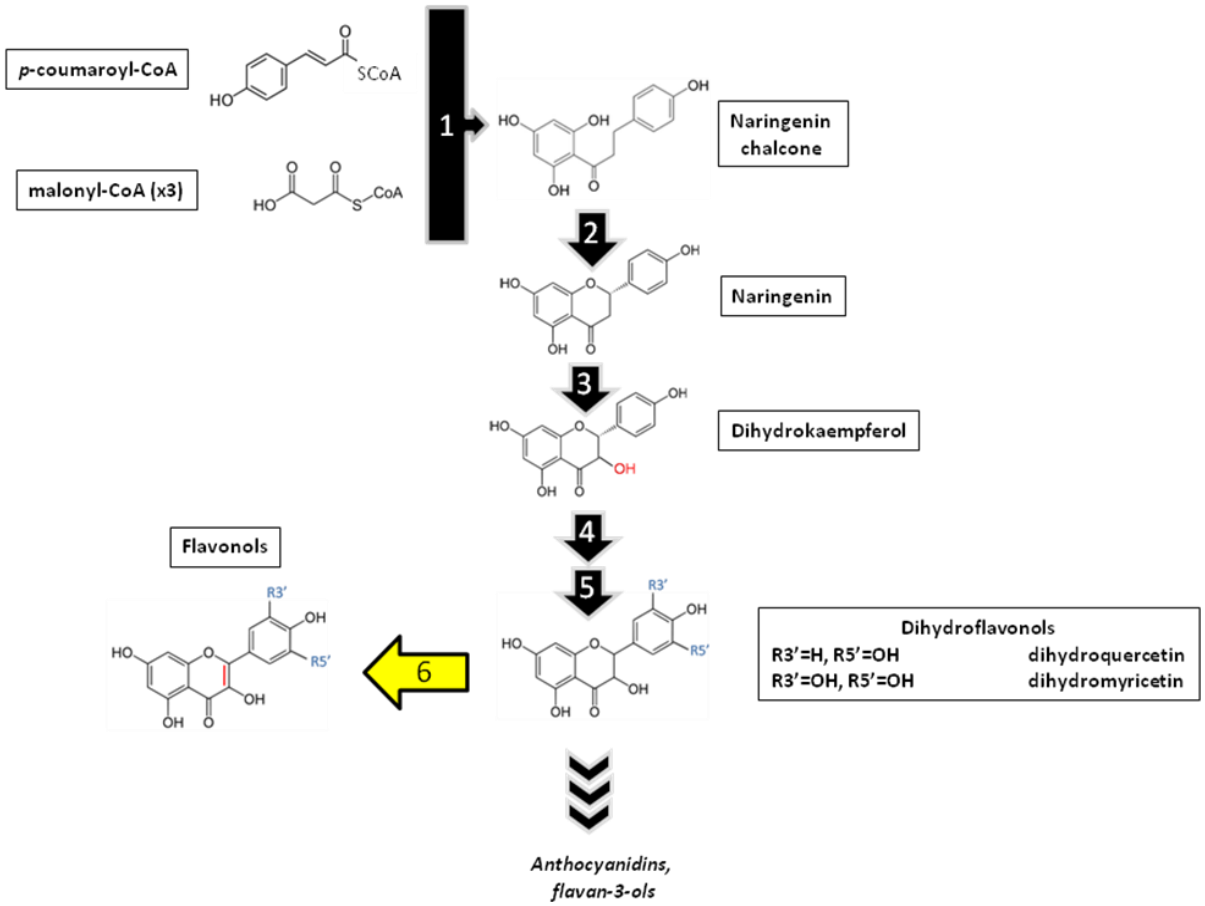


Figure 16. Upper part of the flavonoid pathway.
Enzymes: 1. Chalcone synthase, 2. Chalcone isomerase, 3. Flavanone 3-hydroxylase, 4. Flavonoid-3'-hydroxylase, 5. Flavonoid-3'-5'-hydroxylase, 6. Flavonol synthase.

3.4. Flavonoid pathway

The isolation and characterization of the genes of the flavonoid pathway have progressed significantly since the 1990s. Four plants have attracted particular attention from researchers: maize (*Zea mays*), snapdragon (*Antirrhinum majus*), petunia (*Petunia hybrida*) and *Arabidopsis thaliana* (Winkel-Shirley, 2001). In these species, mutants impaired in the biosynthesis of flavonoids, in particular anthocyanins and PAs, present striking phenotypes (variation in flower and seed colour, respectively).
In this way, the work done in *Arabidopsis thaliana* during the 2000s has allowed the identification of genes involved in the structure and the regulation of the pathway through the study of mutants which present partially or totally discoloured seeds. In these mutants, the accumulation of PAs in the seed endothelium is affected. The corresponding loci have been named transparent testa (tt, Koornneef, 1990). *tt* genes encode TFs, proteins involved in the transport process of flavan-3-ols, or encode a biosynthetic enzyme.
The works performed in the first studied plants have been transposed to more recently sequenced species and/or species of economic interest. In particular, the identification of orthologous genes helped to decipher the flavonoid pathway in grapevine. The key steps of the flavonoid pathway in grapevine are well characterized today, although some particular steps remain to be elucidated.
Indeed, the flavonoid pathway includes a common set of enzymes involved in the common trunk of the flavonoid pathway. However, there is variability in their ability to produce certain flavonoids (e.g. phlobaphenes in maize and sorghum or galloylated PA in tea, persimmon and grape), implying the involvement of specific enzymes.
The enzymes of the flavonoid pathway, like those of the phenylpropanoid pathway, are grouped in metabolons (Wagner & Hrazdina, 1984, Hrazdina *et al*., 1987, Burbulis & Winkel-Shirley, 1999). Some enzymes, belonging to the cytochrome P450 family, are anchored to the membrane of the endoplasmic reticulum (Winkel, 2004).

3.4.1. Enzymatic steps

4-coumaroyl-CoA is the substrate for the first enzyme of the common trunk of the flavonoid pathway, chalcone synthase (CHS). CHS catalyzes the condensation of 4-coumaroyl-CoA with 3 molecules of malonyl-CoA to form naringenin chalcone (Kreuzaler & Hahlbrock, 1972, Fig. 16, step 1). Hence, this enzyme produces the carbon structure (C6-C3-C6), specific to flavonoid. CHS is in competition with stilbene synthase (STS) which produces resveratrol and directs the carbon flux towards the stilbene pathway. STS results from an evolution from CHS and few amino acids are responsible for this switch of activity (Braidot *et al*., 2008). Three *CHS* isogenes have been identified in grapevine (Goto-Yamamoto *et al*., 2002).

Chalcone isomerase (CHI) converts naringenin chalcone into a flavanone, naringenin (Moustafa & Wong, 1967, Fig. 16, step 2).

The following enzyme, flavanone 3-hydroxylase (F3H), hydroxylates the C ring of the flavanone to produce a dihydroflavonol, dihydrokaempferol (Forkmann *et al*., 1980, Fig. 16, step 3). It consists in the last step common to flavonol, anthocyanin and flavan-3-ol biosynthesis.

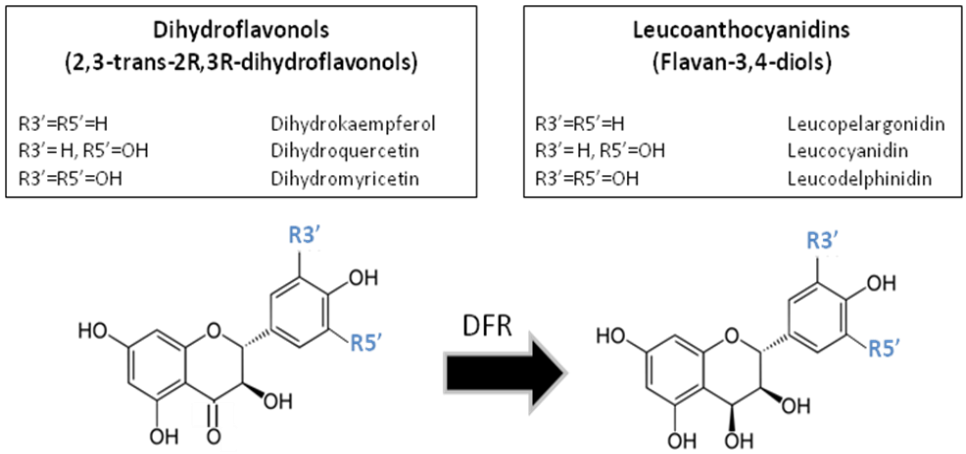


Figure 17. Enzymatic conversion of dihydroflavonols into leucoanthocyanidins catalysed by DFR.

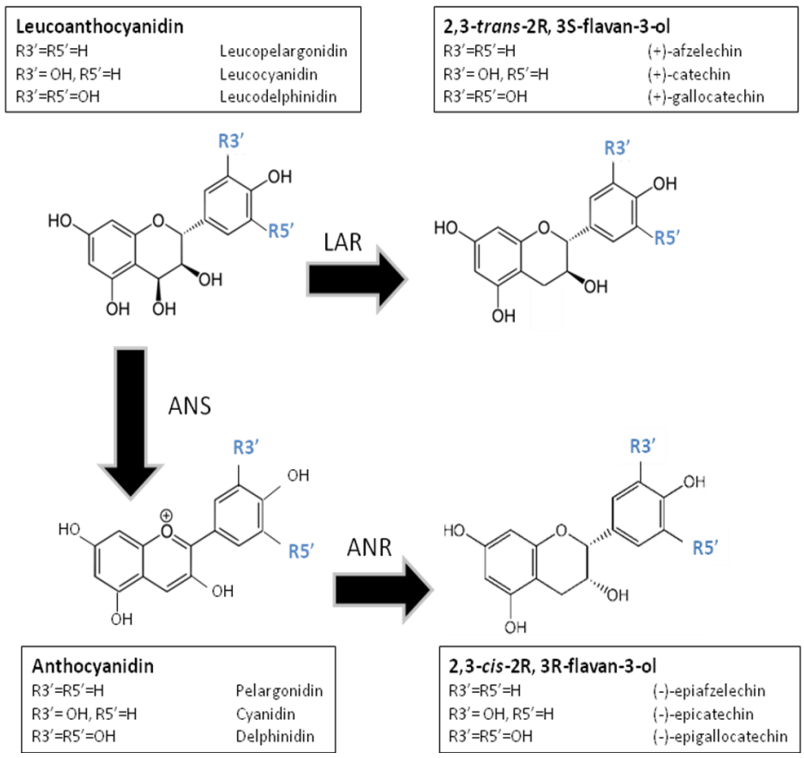


Figure 18. Biosynthesis of flavan-3-ol monomers in the final steps of flavonoid pathway.
ANS: Anthocyanidin synthase, LAR: Leucoanthocyanidin reductase, ANR: Anthocyanidin reductase.

Dihydrokaempferol, which is the dihydroflavonol formed, can be hydroxylated at the 3' position on the B-ring by the flavonoid-3'-hydroxylase (F3'H) to form dihydroquercetin (Froemel *et al*., 1985, Fig. 16, step 4). Grapevine possesses 2 genes encoding F3'H (Castellarin *et al*., 2006).
The hydroxylation can be carried out both at the carbons 3' and 5' *via* the flavonoid-3'-5'-hydroxylase which produces dihydromyricetin (F3'5'H, Menting *et al*., 1994, Fig. 16, step 5). Sixteen *F3'5'H* isogenes have been identified in grapevine genome (Falginella *et al*., 2010).
Both F3'H and F3'5'H of grapevine were functionally characterized in a petunia mutant (Bogs *et al*., 2006).

Flavonol synthase (FLS) is responsible for the dihydroflavonol oxidation to produce flavonols (Holton *et al*., 1993, Fig. 16, step 6). In grape berry, *VvFLS* exhibits a peak of expression in early stage of development and after *véraison* (Fang *et al*., 2013).

Dihydroflavonols can be routed to the biosynthesis of flavan-3-ols or anthocyanins *via* leucoanthocyanidins by dihydroflavonol reductase (DFR, Stafford & Lester, 1982, Fig. 17). Leucopelargonidin, leucocyanidin and leucodelphinidin are the leucoanthocyanidins produced from dihydrokaempferol, dihydroquercetin and dihydromyricetin, respectively. In grapevine, the *DFR* gene has been characterized *in vitro* (Petit *et al*., 2007).

Leucoanthocyanidin dioxygenase (LDOX), also called anthocyanin synthase (ANS), produces anthocyanidins from leucoanthocyanidins (Abrahams *et al*., 2003, Fig 18). LDOX catalyzes the oxidation of leucocyanidin, leucopelargonidin and leucodelphinidin to form respectively cyanidin (magenta color), pelargonidin (orange) and delphinidin (purple). Anthocyanidins are very unstable and must rapidly be glucosylated to form anthocyanins or converted into flavan-3-ols.
Enzymes responsible for anthocyanidin glucosylation have been reported from different plants (see Flavonoid decoration). In grapevine, anthocyanidins are glucosylated by UDP-glucose: flavonoid 3-*O*-glucosyl transferase (UFGT) to form the corresponding 3-glucosides (Ford *et al*., 1998). *UFGT* is not or very few expressed in white cultivars, which explains why anthocyanins are ony detected as traces (Boss *et al*., 1996). In red cultivars, its expression starts at *véraison*.
Flavan-3-ols can be formed from leucoanthocyanidins or anthocyanidins. Two enzymes, each having their substrate specificity, are thereby involved.
Leucoanthocyanidin reductase (LAR) converts the 2,3-*trans*-2R,3S-leucoanthocyanidins into 2,3-*trans*-2R,3S-flavan-3-ols (Fig. 18). *LAR* gene was first cloned from the tropical forage *Desmodium uncinatum* (Tanner *et al*., 2003).
Anthocyanidin reductase (ANR) converts anthocyanidins into 2,3-*cis*-2R,3R-flavan-3-ols in presence of NADPH (Fig. 18). The functional characterization of ANR from *Arabidopsis thaliana*, encoded by the *BANYULS* gene (*BAN*, Devic *et al*., 1999), and *Medicago truncatula* was published in 2003 (Xie *et al*., 2003). *BAN* is only expressed in immature seeds of Arabidopsis whereas *MtANR* is also expressed in flowers and leaves. LAR and ANR have been localized in the cytosol (Pang *et al*., 2007). Grapevine genome carries 2 *LAR* isogenes (*VvLAR1* and *-2*) and one *ANR* gene (*VvANR*). *VvANR* is expressed during early development of berry, in skin and seeds (Bogs *et al*., 2005). The 2 *VvLARs* are expressed in immature seeds. *VvLAR2* also exhibits a peak of expression in skin at *véraison.*

Table 7. Enzymes involved in the flavonoid pathway in grapevine.

Enzyme	Abb.	Nomenclature EC	Grapevine isogenes (NCBI number)	Reference
Chalcone synthase	CHS	EC 2.3.1.74	X75969	Sparvoli *et al*., 1994
			Chs1 (AB015872)	Jeong *et al*., 2004
			Chs2 (AB066275)	
			Chs3 (AB066274)	
Chalcone isomerase *	CHI	EC 5.5.1.6	*Chi1* (X75963)	Sparvoli *et al*., 1994
			Chi2 (TC51784)	Jeong *et al*., 2004
Flavanone 3-hydroxylase	F3H	EC 1.14.11.9	*F3h1* (X75965)	Sparvoli *et al*., 1994
			F3h2	Jeong *et al*., 2004
Flavonoid 3'hydroxylase *	F3'H	EC 1.14.13.21	*F3'h1* (AB213602)	Jeong *et al*., 2006
			F3'h2 (AB213603)	
			F3'h3 (AB213604)	
			F3'h4 (AB213605)	
			VvF3'H (AJ880357)	Bogs *et al*., 2006
Flavonoid 3'-5'hydroxylase	F3'5'H	EC 1.14.13.88	*F3'5'h* (AB213606)	Jeong *et al*., 2006
			VvF3'5'H1 (AJ880356)	Bogs *et al*., 2006
			VvF3'5'H-1a (DQ298201)	Castellarin *et al*., 2006
			VvF3'5'H-2b (DQ298205)	
			F3'5'Ha-p	Falginella *et al*., 2010
Dihydroflavonol reductase *	DFR	EC 1.1.1.219	X75964	Sparvoli *et al*., 1994
Anthocyanin synthase * (Leucoanthocyanin dioxygenase)	ANS (LDOX)	EC 1.14.11.19	X75966	Sparvoli *et al*., 1994
UDP-glucose:flavonoid 3-*O*-glucosyl transferase *	UFGT		X75968	Sparvoli *et al*., 1994
Leucoanthocyanidin reductase	LAR	EC 1.17.1.3	*VvLAR1-1* (AJ865336)	Bogs *et al*., 2005
			VvLAR1-2 (AJ865335)	
			VvLAR2 (AJ865334)	
Anthocyanidin reductase *	ANR	EC 1.3.1.77	*VvANR* (CAD91911)	Bogs *et al*., 2005

*: a single gene copy was found in the grapevine genome (Velasco *et al*., 2007).

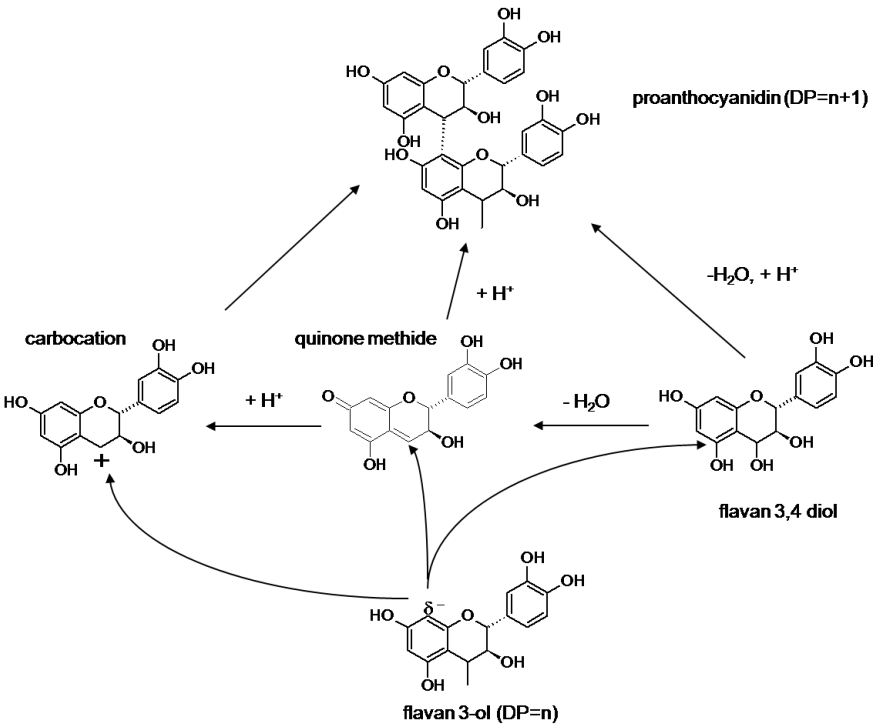


Figure 19. Polymerisation mechanism from flavan-3,4-diol as upper unit.
This figure has been published by Terrier and co-workers, 2009b.

3.4.2. Flavan-3-ol polymerisation

The way PAs are condensed is debated in the literature, the major question being whether this reaction is enzyme dependent or independent (reviewed by Dixon *et al*., 2005, Zhao *et al*., 2010).
Briefly, a precursor molecule (e.g. flavan 3,4 diol, anthocyanidin, epicatechin) must be activated to become electrophilic. Then it undergoes a nucleophilic attack on C4 by a flavan-3-ol and become an upper unit in a PA chain. The most commonly accepted reaction, from flavan-3,4-diol as precursor, is presented in Figure 19.
An unknown enzyme (Polyphenol oxidase (PPO), laccase) could be responsible for this activation. Ectopic overexpression of *MtANR* or *VvANR* in tobacco flowers triggered the production of PA oligomers (Xie *et al*., 2003, Bogs *et al*., 2005). Flav-3-en-3-ol could be intermediate molecule during the enzymatic conversion of anthocyanidin to flavan-3-ol by ANR (Xie *et al*., 2004). However, *in vitro* assay with ANR did not yield PAs. Hence, ANR could provide building blocks for PA construction but must not be involved in PA polymerization.
The PA-deficient *tt10* mutant of Arabidopsis exhibit colourless seeds. *AtTT10* encodes a laccase-like polyphenol oxidase that could be involved in flavan-3-ol oxidation and/or polymerisation (Pourcel *et al*. 2005, Pang *et al*., 2008). However, this mutant had a high content of extractible PAs, suggesting that TT10 is likely involved in the convertion of colourless PAs to brown oxidized PAs, reducing their extractability. *TT10* expression pattern was synchronous with PA accumulation in the vacuole of immature seeds (Kitamura *et al*., 2004). However, *TT10* encodes a protein putatively addressed to the apoplast.

3.4.3. Flavonoid decoration

Flavonoid decoration increases the diversity of those molecules and modifies their properties. Carbon backbone is frequently decorated by –methyl, –glycosyl and –acyl moieties. The sugar or acyl group transferred on flavonoids after their biosynthesis could serve as a tag which eases their transport by tonoplastic transporters and limit their dispersion within the cell. For example, glycosylation could increase anthocyanin stability (Borkowski *et al*., 2005). Methylation can modulate stability, water solubility and color of anthocyanins, that presents a particular interest for wine coloration. An anthocyanin methyltransferase has been characterized in grapevine (Hugueney *et al*., 2009).

The impact of acylation on flavonoids properties and trafficking is described in manuscript #1.

The high diversity of flavonoids due to the different substitutions indicates that numerous non characterized decoration enzymes exist in plant kingdom.

3.4.3.1. Glycosylation

Flavonoid glucosylation is catalysed by UDP-dependant glucosyltransferases (UGTs). Those enzymes often have specificity for the transferred sugar moiety and the position of the carbon of flavonoid backbone as illustrated by the following examples.
In Arabidopsis, UGT79B1 encodes an anthocyanin 3-*O*-glucoside: 2”-*O*-xylosyltransferase (Yonekura-Sakakibara *et al*., 2012). Two glucosyltransferases (At5g17050 and At4g14090), induced

by the anthocyanin-specific MYB TF PAP1, encode flavonoid 3-*O*-glucosyltransferase and anthocyanin 5-*O*-glucosyltransferase, respectively (Tohge *et al*., 2005).

Flavonols and anthocyanins are mainly found as glucosides in grape berry. Two glycosyltransferases called VvGT5 and -6 are involved in the biosynthesis of flavonol 3-*O*-glucuronide and flavonol 3-*O*-glucoside/galactoside, respectively (Ono *et al*., 2010). In grapevine (*Vitis vinifera*), anthocyanins glucosylation is restricted to the activity of UFGT which transfer a glucosyl moiety on the carbon 3 (Ford *et al*., 1998).

In contrast, flavan-3-ol glucosides are minor compounds recently identified in seeds (Delcambre & Saucier, 2012). In *Medicago truncatula*, UGT72L1 can glucosylate epicatechin to form epicatechin 3'-*O*-glucoside (Pang *et al*., 2008). This molecule is transported by the tonoplastic MtMATE1, an ortholog of Arabidopsis TT12, with higher affinity than cyanidin 3-*O*-glucoside (Zhao & Dixon, 2009). This flavan-3-ol glucoside could also be a precursor for PA polymerization.

3.4.3.2. Acylation

In plants, two acyltransferases families are involved in flavonoid acylation.
Acyl moieties can be transferred from acyl-CoA thioesters by BAHD-type acyltransferases (takes its name from the first biochemically characterized enzymes: benzylalcohol *O*-acetyltransferase (BEAT) (Dudareva *et al*., 1998), anthocyanin *O*-hydroxycinnamoyltransferase (AHCT) (Fujiwara *et al*., 1998), anthranilate N-hydroxycinnamoyl/benzoyltransferase (HCBT) (Yang *et al*., 1997) and deacetylvindoline 4-*O*-acetyltransferase (DAT) (St Pierre *et al*., 1998)), or from acyl-glucose esters by Serine CarboxyPeptidases Like-type acyltransferases (SCPL-ATs, Noda, unpublished, Fraser *et al*., 2007).

The sugar moiety of anthocyanins can be acylated with hydroxycinnamates (caffeoyl and *p*-coumaroyl moieties) or acetyl moiety in grapevine. In a recent work, a BAHD-AT able to transfer both caffeoyl, *p*-coumaroyl and acetyl moieties from their respective CoA thioesters to different anthocyanins, has been identified in grapevine (VvAT3, Rinaldo *et al*., 2015). Its poor affinity for malonyl-CoA could explain the lack of malonylated anthocyanins in grapevine.

ATs have been localized in 2 different compartments: cytosol for BAHD-ATs and vacuole for SCPL-ATs. The comparison of the two enzymatic systems involved in acylation of phenolic compounds has been reviewed in Bontpart *et al*. (2015), see Chapter 4, manuscript #2.

Flavan-3-ol monomers or PA subunits can be acylated with a galloyl moiety. The most successful study concerning flavan-3-ol galloylation has been conducted from tea leaves and published recently (Liu *et al*., 2012). Galloylated flavan-3-ols (ECG and EGCG) were produced from β-G and flavan-3-ols (EC and EGC) as substrates with an enzyme extract from tea leaves. The authors propounded that this reaction could be catalysed by SCPL-AT(s) *in planta*. However, the gene(s) involved in flavan-3-ol galloylation in plants are still unknown.

3 glucosyltransferases (VvGTs) induced by MybPAs TFs have already been characterized in the laboratory (Khater *et al*., 2012). These enzymes have the ability to glucosylate different HBAs and

HCAs, and notably GA to form β-G, which could be the precursor (galloyl donnor) of galloylated PAs.

Otherwise, HCA glucose esters produced by those enzymes could also be precursors for SCPL-AT catalyzed acylation of anthocyanins and/or tartrate. The data collected in the literature indicate that anthocyanins acylation in grapevine can be catalysed both by BAHD-AT(s) and SCPL-AT(s). However, few data are available concerning the biosynthesis of hydroxycinnamoyl tartrates and the gene(s) responsible for the biosynthesis of those metabolites have not been identified.

The coexpression of two putative SCPL-ATs (VvGATs) with VvGTs in grapevine hairy-roots overexpressing MybPA TFs (Terrier *et al*., 2009a) lets think that VvGATs could use those phenolic glucose esters to catalyse flavan-3-ols galloylation and/or the biosynthesis of hydroxycinnamoyl tartrates.

4. Transport and storage of flavonoids

By analogy with xenobiotics, flavonoids can be considered as potentially toxic if they disperse in the cell. Thus, different mechanisms are used for their transport to their storage compartment, where they reach sufficient concentration to ensure their biological roles with no harmful effect on other cell mechanisms.
Once synthesized at the cytosolic side of the endoplasmic reticulum (Zhao & Dixon, 2010), flavonoids are expected to be transported and stored mainly in the vacuole (Kitamura, 2006). However, other sites were described as flavonoids storage sites: cell wall, chloroplast, nucleus, cytoplasm (reviewed by Agati *et al*., 2012).

Microscopic observations have located the anthocyanins in the vacuole of the cells composing the outer layers of the skin hypodermis of the red grape berry. In the grape berry, PAs are located both in the vacuole and cell wall (Gény *et al*., 2003, Downey *et al*., 2003b, Gagné *et al*., 2006).
Although flavonoids biosynthetic mechanisms are widely characterized, those relating to their intracellular transport remain relatively limited to this day. Studies on the transport of flavonoids demonstrate the existence of a network of vesicles (vesicular transport) as well as membrane transporters (Fig. 20).

4.1. Proteins involved in flavonoids transport

4.1.1. Glutathione S-transferase (GST)

Glutathione S-transferases (GSTs) would have the ability to combine a tripeptide (glutathione) to molecules to facilitate their transport through ABC (ATP-binding cassette) GS-X pumps. The existence of such a mechanism for the transport of flavonoids has been considered. GSTs involved in flavonoid transport have been identified in various plants: BZ2 in maize (Marrs *et al*., 1995), AN9 in petunia (Alfenito *et al*., 1998) and TT19 in *A. thaliana* (Kitamura *et al*., 2004). The mutant *tt19*

accumulate less anthocyanins in vegetative organs and less PAs in seeds, suggesting that TT19 would be involved in the accumulation of both anthocyanins and PAs (Kitamura *et al*., 2004).
However, the lack of glutathione-linked flavonoids suggests that GSTs would rather function as a ligand capable of escorting the flavonoids from cytosol to vacuolar transporters (Mueller & Walbot, 2001). The flavonoid-GST complex would prevent the oxidation of flavonoids (Kitamura *et al*., 2010).
In grapevine, a GST was localized in the cytoplasm and is involved in the storage of anthocyanins directly in the vacuole rather than in the filling of vesicles (Gomez *et al*., 2011).
This GST is related to accumulation of anthocyanins and can complement the maize mutant *bz2* (Conn *et al*., 2008). To date, no GST that would be involved in PA accumulation has been identified yet in grapevine (Terrier *et al*., 2009a).

4.1.2. ATP-binding cassette (ABC)-type transporters

Flavonoid transport can be achieved by ABC-type primary transporters which are composed of an ATP binding domain, and couple the hydrolysis of ATP to the passage of molecules, independently of pre-established H^+ gradient. These carriers have been particularly associated with the transport of xenobiotics and flavonoid glycosides (Frangne *et al*., 2002, Goodman *et al*., 2004). In maize, a Multidrug Resistance associated Protein-carrier (MRP) which belongs to a subgroup of the ABC transporters family has been identified (Goodman *et al*., 2004). This protein (ZmMRP3) was localized on the tonoplast. Mutants affected for this gene exhibit a reduction in the accumulation of anthocyanins.
The ABC-type proteins found in eukaryotic organisms can be classified into 8 groups (Verrier *et al*., 2008). The grapevine genome possesses 135 open reading frames (ORFs) of ABC-type proteins, 120 encoding membrane proteins and 15 lacking a transmembrane domain (Çakır & Kılıçkaya, 2013).
In grapevine, an ABC-type protein (ABCC1) allows the transport of anthocyanin 3-*O*-glucosides in the vacuole and was detected at the tonoplast (Francisco *et al*., 2013). This protein is expressed in exocarp with a peak at *véraison* that is synchronous with anthocyanin accumulation in grape berry. No ABC-type protein involved in flavan-3-ols transport has been identified yet in grapevine.

4.1.3. MATE-type transporters

MATE (Multidrug And Toxic Compound Extrusion)-type proteins are secondary transporters which take advantage of the electrochemical gradient generated by proton pumps.
MATE-type transporters are ubiquitous proteins found in prokaryotes and eukaryotes (Moriyama *et al*., 2008).
A MATE-type transporter was identified from the *A. thaliana tt12* mutant, which does not accumulate PAs in the endothelium of seeds (Debeaujon *et al*., 2001). TT12 protein fused to GFP (Green Fluroescent Protein) has been localized at the tonoplast (Marinova *et al*., 2007). However, yeast vesicles expressing TT12 can carry cyanidin 3-*O*-glucoside but not catechin 3-*O*-glucoside, despite its presumed role in PA accumulation. In *M. truncatula*, MATE1 is an ortholog of TT12. This carrier transports cyanidin 3-*O*-glucoside but has a greater affinity for epicatechin 3'-*O*-glucoside (Zhao & Dixon, 2009).
In grapevine, 2 MATE-type transporters involved in the transport of anthocyanins, called anthoMATE1 and -3 (AM1 and -3) have been identified (Gomez *et al*., 2009). *In vitro*, yeast vesicles expressing these anthoMATEs specifically carry acylated anthocyanins. The overexpression of AM-

GFP construction in grape hairy-roots helped locate these proteins at the tonoplast and on membranes surrounding cytoplasmic vesicles.
MATE-type proteins (VvMATE1 and -2) are suspected to be involved in the transport of flavan-3-ols in grapevine (Pérez-Díaz *et al*., 2014). Indeed, *VvMATE1* and *-2* are mainly expressed in immature seeds, which coincides with the accumulation of PAs in this tissue. The expression of these proteins fused to GFP in Arabidopsis protoplasts allowed to localizeVvMATE1 at the tonoplast and VvMATE2 at the Golgi apparatus. However, their precise function including substrate specificity was not investigated.

4.1.4. ATPase proton pump

Vacuolar pH is important for the regulation of flavonoid transport. The activity of the MATE-type transporters is closely linked to the pH gradient generated by ATPase proton pump.
In plants, 3 types of proton pumps are involved in pH regulation. P-type H^+-ATPases are usually localized in the plasma membrane. Otherwise, vacuolar-type H^+-ATPases (V-ATPases) and vacuolar H^+-pyrophosphatases (H^+-PPases) are found on endomembranes.
However, the P-type ATPase PH5 from petunia has been localized on the tonoplast (Verweij *et al*., 2008). *Ph5* mutant has lost its flower colour due to a lack of vacuolar acidification and accumulated less PAs in seeds. In Arabidopsis, H^+-ATPase isoform 10 (AHA10) belongs to the P-type ATPase family. It was mainly expressed in immature seeds and responsible for PAs accumulation (Baxter *et al*., 2005). Indeed, the *tt* mutant *aha10* had 100 fold less PAs in seeds. More recently, it has been shown that AHA10 was localized on the tonoplast (Appelhaggen *et al*., 2015).
A grapevine P-type ATPase, putatively involved in flavan-3-ols transport is studied in our lab.

4.1.5. BTL-like transporter

A transporter homologue to mammalian bilitranslocase (BTL) was for the first time in plants detected in carnation petals using antibodies (Passamonti *et al*., 2005a). In mammals, BTL is involved in the transport of bilirubin, a yellow pigment resulting from hemoglobin degradation, from blood plasma to liver (Passamonti *et al*., 2005b). Liver BTL also exhibits affinity for flavonoids (Karawajczyk *et al*. 2007).
The involvement of this kind of transporter in flavonoids transport using similar antibodies has been investigated in red (Braidot *et al*., 2008) and white grape berries (Bertolini *et al*., 2009). In these two studies, BTL-like transporter was detected in hypodermal layers (skin) and vascular bundles (pulp) but at different developmental stages. BTL-like transporter could be involved in the transport of flavonols in white cultivars besides anthocyanins in red cultivars.
However, the putative gene(s) encoding grape BTL-like transporters have not been characterized in grapevine yet.

Table 8. Examples of proteins involved in flavonoid transport in plants.

Type	Protein	Transported flavonoids	Species	Reference
GST	TT19	Anthocyanins / Flavan-3-ols	*Arabidopsis thaliana*	Kitamura *et al.*, 2004
	BZ2	Anthocyanins	*Zea mays*	Marrs *et al.*, 1995
	AN9	Anthocyanins	*Petunia hybrida*	Alfenito *et al.*, 1998
	GST4	Anthocyanins	*Vitis vinifera*	Ageorges *et al.*, 2006; Conn *et al.*, 2008
ABC	ABCC1	Anthocyanins	*Vitis vinifera*	Francisco *et al.*, 2013
	MRP	Anthocyanins	*Zea mays*	Goodman *et al.*, 2004
MATE	VvAM1 / -3	Anthocyanins	*Vitis vinifera*	Gomez *et al.*, 2011
	VvMATE1 / -2 *	Flavan-3-ols	*Vitis vinifera*	Perez-Diaz *et al.*, 2014
	TT12	Flavan-3-ols	*Arabidopsis thaliana*	Debeaujon *et al.*, 2001; Marinova *et al.*, 2007
P-ATPase	PH5	Anthocyanins	*Petunia hybrida*	Verweij *et al.*, 2008
	AHA10/TT13	Flavan-3-ols	*Arabidopsis thaliana*	Baxter *et al.*, 2005 Appelhaggen et al., 2015
	VvP-ATAse *	Flavan-3-ols	*Vitis vinifera*	unpublished
BTL-like	BTL *	Anthocyanins/ Flavonols	*Vitis vinifera*	Braidot *et al.*, 2008 Bertolini *et al.*, 2009

*: Not fully characterized.

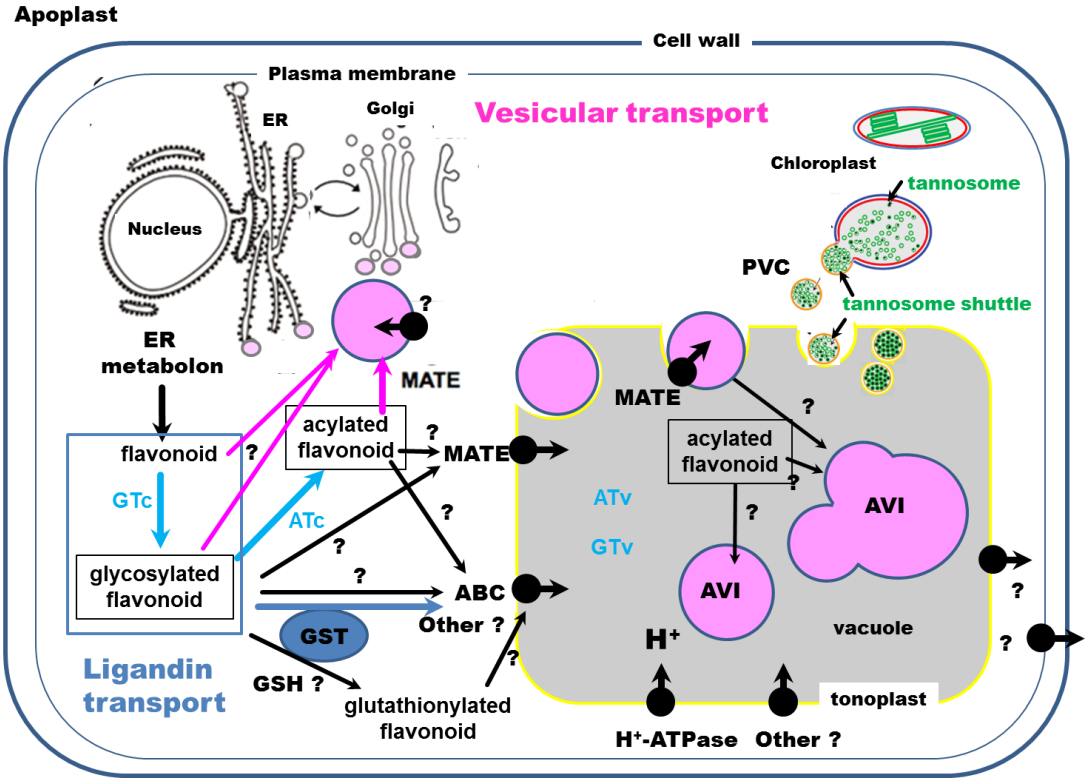


Figure 20. Models for flavonoids transport in a plant cell of grapevine.
Abbreviations: ER: endoplasmic reticulum, TGN: trans Golgi network, PVC: prevacuolar compartment, AVI: Anthocyanic vacuolar inclusions, MATE: Multidrug And Toxic Compound Extrusion, ABC: ATP-binding cassette, GST: Glutathione *S*-transferase, GSH: Glutathion, GTc: cytosolic glucosyltransferase, GTv: vacuolar glucosyltransferase, ATc: cytosolic acyltransferase, ATv: vacuolar acyltransferase.

4.1.6. Vesicular transport model

The hypothesis of a vesicular network to route flavonoids from their biosynthesis site (endoplasmic reticulum) to their storage site, involving the Golgi apparatus, has been proposed (Grotewold, 2004, 2006). These vesicles could achieve high levels of anthocyanins in the vacuole and thus intensify and stabilize the colour of the coloured organs (e.g. flowers, berries).
Anthocyanins have been localized in cytoplasmic vesicles called anthocyanoplasts (ACPs, Calderón *et al*., 1993). ACPs would be a transport mechanism facilitating the delivery of anthocyanins to the vacuole.
Such vesicles derived from endoplasmic reticulum could be fused to the vacuole via autophagy (Pourcel *et al*., 2010).
Some vesicles, rich in anthocyanins, were called anthocyanic vacuolar inclusions (AVIs) and have been observed in the vacuole in lisianthus (*Eustoma grandiflorum*, Markham *et al*., 2000, Zhang *et al*. 2006), grapevine (Conn *et al*., 2003), and *Arabidopsis thaliana* (Poustka *et al*., 2007).The overexpression of the TF MybA1 in grapevine hairy-roots allowed to observe AVIs (Cutanda-Perez *et al*., 2009, Gomez *et al*., 2011). In grape cells suspensions, an exogenous treatment with sucrose, jasmonic acid and light triggered an increase of the AVI number (Conn *et al*., 2010).

Abrahams *et al*. (2002) proposed a PA transport model through cytoplasmic vesicles that can fuse with the central vacuole. More recently, the chloroplast has been proposed as an organelle where could take place the biosynthesis and polymerisation of PAs in young tissues (Brillouet *et al*., 2013). Tannin accretions have been observed in the thylakoidal lumen. TEM (transmission electron microscopy) observations suggest that thylakoids pearl to form vesicles called tannosomes which are then encapsuled in chloroplast envelopes to form shuttles of tannosomes finally exported and fusionned to the vacuole. Those shuttles could be incorporated by an autophagy-like mechanism (Kulich & Žárský, 2014).
The immunolocalization of phenylpropanoid/flavonoid pathway enzymes (C4H, Chen *et al*., 2006, CHS, Tian *et al*., 2008, ANS, Wang *et al*., 2010) in the chloroplast of immature grape berries supports that flavonoids biosynthesis could occur this organelle. Otherwise, flavonoids have been localized in the chloroplast (e.g. catechin, Liu *et al*., 2009). However, evidences for a complete pathway in this organelle are lacking.

5. Transcription factors regulating the flavonoid pathway

TFs orchestrate secondary metabolites biosynthetic pathways in plants by modulating the expression of biosynthetic enzymes. More specifically, TFs are proteins able to interact physically with DNA sequences called *cis* sequences and modulate the target gene(s) expression. Such *cis* sequences are usually located within the gene promoter upstream of the coding part of the target gene. TFs can promote or inhibit the expression of a gene.
Isolation and characterization of genes encoding TFs have revealed how the enzymes that constitute the biosynthetic pathways are regulated. TFs are "molecular clocks" that can control the biosynthesis pathway in the different developmental stages and plant tissues. They can be regulated by external signals, consequence of the (a)biotic interactions of the plant with its environment, or regulated by upstream TFs. TFs are also biotechnological tools used to modulate the accumulation of compounds *in*

planta (Falcone Ferreyra *et al*., 2012).TFs controlling the flavonoid pathway are certainly the most studied and best characterized TFs in the plant kingdom (reviewed by Lepiniec *et al*., 2006, Liu *et al*., 2015a). Flavonoid pathway regulation is a complex mechanism which involves a regulatory loop between regulators (Dubos *et al*., 2008; Xu *et al*., 2013).

Two main TF families are involved in the regulation of the flavonoid pathway (Koes *et al*., 2005): bHLH (basic helix-loop-helix) and R2R3-MYB (V-myb myeloblastosis viral oncogene homolog). By associating with WDR (tryptophan-aspartic acid (W-D) dipeptide repeat), they form the MYB-bHLH-WD40 complex (MBW) that regulates the expression of genes of the flavonoid pathway by interacting directly with DNA (reviewed by Hichri *et al*., 2011). R2R3-MYB and bHLH TFs regulate gene expression interacting directly with *cis* elements whereas WDR would be involved in the interaction with both TFs (Baudry *et al*., 2004). The best characterized MBW complex is certainly the one that governs the biosynthesis of PAs in *Arabidopsis thaliana*. Composed of the proteins TRANSPARENT TESTA2 (TT2, R2-R3 MYB), TRANSPARENT TESTA8 (TT8, bHLH) and TRANSPARENT TESTA GLABRA1 (TTG1, WDR), it induces the expression of *DFR*, *LDOX* and *ANR* (Nesi *et al*., 2001, Debeaujon *et al*., 2003, Baudry *et al*., 2004). More recently, it has been shown that this MBW complex also regulates proteins involved in flavonoid transport: the Glutathione *S*-transferase TT19, the MATE-type transporter TT12 and the P-ATPase AHA10 (Xu *et al*., 2014).

Myb TFs ensure the regulation specificity. The most studied subgroup (R2R3-Myb) of flavonoid TFs in plants will be briefly described and we will focus then on discovered R2R3-Myb involved in the control of grapevine flavonoids.

5.1. R2R3-MYB TFs in plants

The MYB proteins belong to one of the most important plant families of TFs and regulate among other cellular processes the phenylpropanoid/flavonoid metabolism (Liu *et al*., 2015a).These proteins have a varying number of DNA binding domains (MYB domain), which allows their separation into 4 classes. The most common class is the R2R3-MYBs with between 100 and 200 members by species (Stracke *et al*., 2001, Yanhui *et al*., 2006, Matus *et al*., 2008, Wilkins *et al*., 2009). R2R3-MYB TFs possess a highly conserved MYB domain in their N-terminal region, constituted of imperfect repeats (R). Each repeat consists of 50 nucleotides encoding 3 α-helices. The activation or repression domain in C-terminal, more divergent, allows phylogenetic analyses.
The first genes encoding R2R3-type MYB FTs, *C1* and *PL* were discovered in maize where they regulate the production of anthocyanins (Cone *et al*., 1986, 1993).
More than 100 members constitute the R2R3-MYB TFs family that regulates positively or negatively the biosynthesis of various branches of the flavonoid pathway (flavones, flavonols, anthocyanins, PAs). The analysis of their sequence allowed their classification into a dozen groups whose members are usually involved in one of the branches of the flavonoid pathway (Dubos *et al*., 2010). However, the TFs of the same group do not necessarily activate the same genes. Several TFs controlling the flavonoids pathway were identified in grapevine and recently reviewed (Czemmel *et al*., 2012).

5.2. Grapevine R2R3-MYB TFs

In addition to their specificity for the biosynthesis of precise flavonoids (anthocyanins, flavonols and PAs), specific TFs coordinately activate genes of the common core of the biosynthetic pathway.
Myb5a and -b regulate the expression of enzymes from the upper part of the flavonoid pathway (Deluc *et al*., 2006, 2008, Table 9). Myb5a is mainly expressed in the skin and seeds at the green stage. Myb5b is mainly expressed during the ripening of the berry (Deluc *et al*., 2008). Overexpression of these FTs in tobacco causes an accumulation of anthocyanins and PAs.

Contrary to the other Myb-TFs involved in flavonoids regulation, flavonol specific Myb-TFs are bHLH independent (Mehrtens *et al*., 2005).
MYBF1 is recognized as a TF specific of the flavonol pathway. Before flowering, it induces the expression of *FLS1* (Czemmel *et al*., 2009). Its expression, like that of *FLS1*, exhibits a peak during flowering and in skin after *véraison*. Furthermore, MYBF1 is regulated by light, independently of the development stage.

The anthocyanin biosynthesis is regulated by 2 TFs: MybA1 and MybA2, from *véraison* till maturation (Kobayashi *et al*., 2002, 2004, Walker *et al*., 2007). Kobayashi and co-workers (2004) have shown that the loss of pigmentation in white berries is linked to a retrotransposon-induced mutation in *VvMybA1*.
MybA1 notably activated the expression of *UFGT*, a *GST* isogene, an *O-methyltransferase* gene and an *AnthoMATE* transporter gene (Cutanda-Perez *et al*., 2009).

PAs biosynthesis is activated by several TFs: MybPA1 (Bogs *et al*., 2005), MybPA2 (Terrier *et al*., 2009a), and MybPAR (Koyama *et al*., 2014) and inhibited by MybC2-L1 (Huang *et al*., 2014) and MybC2-L3 (Cavallini *et al*., 2015).

MybPA1 complements the Arabidopsis PAs-deficient *tt2* mutant.
Another TT2-like Myb TF, called MybPA2 (Myb Proanthocyanidins 2) by analogy with MybPA1 (Bogs *et al*., 2007) has been characterized in the lab (Terrier *et al*., 2009a). The expression profile of MybPA1 and MybPA2 along berry development suggests that these TFs control the PAs pathway in a tissue-specific manner (Terrier *et al*., 2009a). Indeed, their expression pattern in seeds and leaves coincides with the accumulation of PAs from flowering to *véraison*.
MybPA-1 and -2 have been overexpressed in grapevine hairy-roots to identify genes potentially involved in PAs biosynthesis. Among the 51 genes induced by both TFs in grapevine hairy-roots, Terrier and coworkers (2009) identified genes from the shikimate pathway (*DAHPS*, *SDH9*, a *SK*), the phenylpropanoid pathway (*PAL* and *4CL*), the flavonoid pathway (*CHS*, 2 isogenes of *F3H*, *F3'H*, *ANR*, *LDOX*, a *TT12-like MATE*). Moreover, putative "decoration" enzymes (VvGT2 and -3, VvGAT1) and a MATE-type transporter were also identified.
MybPAR was mainly expressed in the seeds around *véraison*, and in the skin of immature berries (Koyama *et al*., 2014).
MybPAR expression was modulated by light exposure and exogenous ABA in skin (Koyama *et al*., 2014). A transient reporter assay monitored in grapevine cells has shown that MybPAR was able to activate the promoter of genes involved in the common flavonoid pathway (*VvCH3*, *VvF3'5'Hd*), the PA-specific pathway (*VvANR*, *VvLAR1*, *VvLAR2*), decoration (*VvGT1* and *VvGAT2*), transport (*TT12-like MYB*), but did not activate UFGT promoter.

Table 9. Enzymes regulation and expression pattern of positive proanthocyanidins R2R3-MYB TFs.
This figure is adapted from Lai *et al*., 2013.

Espèce	**FT**	**Induced genes**	**Organ/tissue, context**
Arabidopsis thaliana	AtTT2	*CHS, CHI, F3H, F3'H, FLS, DFR, ANS, BAN TT19, TT12, AHA10*	Seeds
Brassica napus	BnMYBTT2	?	Seeds
Diospyros kaki (Persimmon)	DkMYB2 DkMYB4	*CHS, F3'H, F3'5'H, DFR, ANR, LAR* *ANR, LAR*	Wounding induced Fruit, Wounding induced
Fragaria x *ananassa* (Strawberry)	FaMYB9 FaMYB11	*F3'H, DFR, ANS, 3GT, ANR, LAR* *F3'H, DFR, ANS, 3GT, ANR, LAR*	Immature fruits Immature fruits
Hordeum vulgare (Barley)	HvMYB10	?	
Lotus japonicus	LjMYBTT2	*ANR*	Roots, stems
Medicago truncatula (Barrel medic)	MtPAR	*WD40*	Seed coat
Populus tremuloides (Poplar)	PtMYB134	*CHS, CHI, F3H, DFR, ANR, ANS, LAR*	Leaves
Prunus persica (Peach)	PpMYBPA1	*DFR, LAR, UFGT*	Fruits
Theobroma cacao (Cacao tree)	TcMYBPA	*CHS, CHI, FLS, F3H, DFR, ANS, BAN, UFGT*	PAs and anthocyanins
Trifolium arvense (Haresfoot clover)	TaMYB14	*F3'5'H, DFR, ANS*	Leaves
Vaccinium corymbosum (Blueberry)	VcMYBPA1	?	Immature berries
Vitis vinifera (Grapevine)	**VvMYBPA1**	*CHI, F3'5'H1, ANS, LAR1, ANR, GATs, MATE*	Mainly in seeds
	VvMYBPA2	*DFR, ANS, LAR1, LAR2, ANR, GT, MybPA1*	Mainly in berry skin
	VvMYB5a	*CHI, F3'5'H, ANS, ANR, LAR1, UFGT*	Berries
	VvMYB5b	*CHI, F3'5'H, ANS, ANR, LAR1, UFGT*	Berries
	VvMybPAR	*CH3, VvF3'5'Hd, ANR, LAR1, LAR2 GTs, GATs*	Berry skin and seeds

References: AtTT2 (Nesi *et al*., 2001; Baudry *et al*., 2004, Xu *et al*., 2014); BnMYBTT2 (Wei *et al*., 2007); DkMYB2 (Akagi *et al*., 2010); DkMYB4 (Akagi *et al*., 2009); FaMYB9 & FaMYB11(Schaart *et al*., 2013); HvMYB10 (Himi *et al*., 2012); LjMYBTT2 (Yoshida *et al*., 2008); MtPAR (Verdier *et al*., 2012); PpMYBPA1 (Ravaglia *et al*., 2013); PtMYB134 (Mellway *et al*., 2009); TcMYBPA (Liu *et al*., 2015b); TaMYB14 (Hancock *et al*., 2012); VcMYBPA1 (Zifkin *et al*., 2012); VvMYBPA1 (Bogs *et al*., 2007; Czemmel *et al*., 2009); VvMYBPA2 (Terrier *et al*., 2009a); VvMYB5a (Deluc *et al*., 2006); VvMYB5b (Deluc *et al*., 2008), MybPAR (Koyama *et al*., 2014).

6. Objectives

The literature review highlights the lack of knowledge about certain steps in the PAs biosynthesis. In particular, the mechanisms involved in polymerization and galloylation of flavan-3-ols and also some trafficking events are not fully understood.

Contrary to *Arabidopsis thaliana*, grapevine genetic mutants are not available and stable transformation of a whole plant results in a long process.
However, numerous large scale studies performed previously in our team helped to find genes putatively involved in PAs biosynthesis: transcriptomic analysis (Terrier *et al*., 2009a), quantitative genetic analysis (Huang *et al*., 2012), association genetic study (Carrier *et al*., 2013).

Cross-data mining led to establish a list of genes involved in PAs biosynthesis (Carrier *et al*., 2013)
This list highlighted genes that encode proteins potentially involved in flavan-3-ol galloylation: a SDH, 3 yet characterized UGTs (VvGT1, -2 and -3, Khater *et al*., 2012) and 2 SCPL-ATs (glucose acyltransferases, VvGAT1 and -2). Those candidate genes were induced by at least MybPA1 or MybPA2 in transgenic hairy-roots, and differentially expressed in skin during berry development. Moreover, those genes were found under QTL intervals influencing PA %G in grape berry skin. Another transcriptomic study confirmed the involved of all these selected genes in PA biosynthesis (Koyama *et al*., 2014).

VvGAT1 was identified as a strong candidate for flavan-3-ol galloylation, and was tested using association genetics. The exonic sequences of VvGAT1 were sequenced using a core collection of 141 individuals to identify a significant effect of sequence polymorphism on the phenotype. One SNP (Single-nucleotide polymorphism) was significantly associated with %G variation.

This list of putative candidates, and the bibliographic study presented in this first chapter, led us to formulate the following hypothesis:
- A SDH protein would be responsible for the biosynthesis of GA from a shikimate pathway intermediate, the 3-dehydroshikimate.
- The already characterized VvGTs are able to catalyse the formation of β-G from GA and UDP-Glucose. They are also able to form hydroxycinnamoyl glucose esters from HCAs and UDP-Glucose. VvGATs would be involved in the transfer of a galloyl moiety from β-G to flavan-3-ols. They could also transfer hydroxycinnamoyl moieties from hydroxycinnamoyl glucose esters to tartaric acid.

To explore these hypotheses, the functional validation of VvSDHs and VvGATs was undertaken by a homologous (grapevine hairy-roots or leaves) and/or heterologous (*Escherichia coli* or tobacco leaves) expression.
The expression profile of these genes during the development of the grape berry and in various tissues was determined.

CHAPTER 2: MATERIAL AND METHODS

1. BIOLOGICAL MATERIAL

1.1. Plant material

1.1.1. Grape berries

The same batch of grape berries cv. Syrah, harvested in summer 2004 on the vineyard of Montpellier SupAgro campus, was used for gene expression studies. The berries were harvested at 11, 18, 35, 52, 56, 58, 65, 86 and 99 daf, frozen in liquid nitrogen and stored until further use. 52 daf corresponds to *véraison* stage. At 18, 52 and 99 daf, additional samples were prepared for which the skin, the pulp and the seeds were separated. The berries were washed, weighted and counted before being frozen in liquid nitrogen. The samples were cold-milled using a Dangoumeau ball mill (Dangoumill 300, Longjumeau, France) and stored at -80°C.

1.1.2. Grapevine seedlings grown *in vitro*

Grapevine stems grown on the Montpellier Supagro campus were removed and disinfected with bleach. After several rinses in sterile water, stem segments of a few centimeters including an internode were cut and transplanted into sterile glass tubes containing solid MS/2 medium (Murashige & Skoog medium, see Annexes). The seedlings were transplanted regularly by cuttings.

1.1.3. Tobacco *Nicotiana benthamiana*

Nicotiana benthamiana tobacco seeds were stored at room temperature in 1.5 mL Eppendorf tubes. A growing pot (7 x 7 x 7 cm) was filled to the brim with organic compost (Neuhaus S) and placed on a tray. At the base of the pot, holes allowed to hydrate the substrate by capillarity after irrigation of the culture tray. When the water had reached visibly the substrate surface, fifteen seeds were sown on the surface of the pot. An upturned plastic pot was placed over the pot to retain maximum moisture. The pot was placed in a culture chamber (temperature: 24°C) and watered every 2-3 days. Ten days after sowing, the seedlings were transplanted using tweezers in individual pots previously irrigated by capillarity.
Up to 24 pots have been randomized per tray and placed in a culture chamber. The trays were irrigated every 3-4 days with 1 L water.
4-6 weeks old plants were used for transient transformation by agroinfiltration.

Plant material was grown in greenhouse (26°C, 15h photoperiod).

1.2. Bacterial strains

The bacterial strains and the culture conditions used in this study are listed in Table 10.

1.2.1. *Escherichia coli* DH5α™.

The strain DH5α™ of *Escherichia coli* (Invitrogen) was used for the incorporation of the cloning vectors of the Gateway® technology. Its genotype is: F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-.

The *gyrA96* mutation allows the incorporation and multiplication of Gateway® technology vectors carrying the *ccdB* gene.

1.2.2. *Escherichia coli* BL21 (DE3).

The BL21 strain of *E. coli* has been used to incorporate the pGEX-4T2 vector for heterologous expression of grapevine SDHs. This strain lacks the *LON* gene encoding a protease and is thus suitable for the production of recombinant proteins. This strain was grown in LB medium (see Annexes) supplemented with the appropriate antibiotic.

1.2.3. *Escherichia coli* ccdB Survival™.

One Shot® ccdB Survival™ 2 T1^{R} Chemically Competent Cells (Invitrogen) were used to propagate and maintain vectors containing the *ccdB* gene (e.g. pEAQ).

1.2.4. *Agrobacterium rhizogenes* A4.

Agrobacterium rhizogenes A4 strain comes from the Centre Français des Bactéries Phytopathogènes d'Angers (http://www-intranet.angers.inra.fr/cfbp/).
This strain was grown in medium MGL/B (see Annexes) supplemented with the appropriate antibiotic.

1.2.5. *Agrobacterium tumefaciens* strains.

The GV3101 strain (provided by Carine Alcon, INRA Montpellier) was grown in LB medium supplemented with rifampicin (25 µg.mL^{-1}). The C58C1 strain (provided by Pere Mestre, INRA Colmar) was grown in LB medium supplemented with tetracycline (5 µg.mL^{-1}) and a second antibiotic depending on the inserted vector.

1.2.6. Storage of microbial strains.

All the microbial strains used in this study were grown during one (*E.coli*) or 2 days (Agrobacteria) in the adequate conditions, mixed with 20% (v/v) glycerol and stored at -80°C.

Table 10. Bacterial strains and culture conditions.

Microorganism	**Strains**	**Culture medium**	**Temperature (°C)**	**Used antibiotic**
Escherichia coli	DH5α	LB	37	
	BL21			
	ccdB Survival			
Agrobacterium rhizogenes	A4	MGL/B	28	
Agrobacterium tumefaciens	GV3101	LB	28	Rifampicin
Agrobacterium tumefaciens	C58C1	LB	28	Tetracyclin

1.3. Vectors

The vectors used in this study are listed in Table 11.

1.3.1. pGEM®-T Easy

The pGEM®-T easy vector (Promega, Madison, USA) was used to clone the cDNA of VvSDHs.

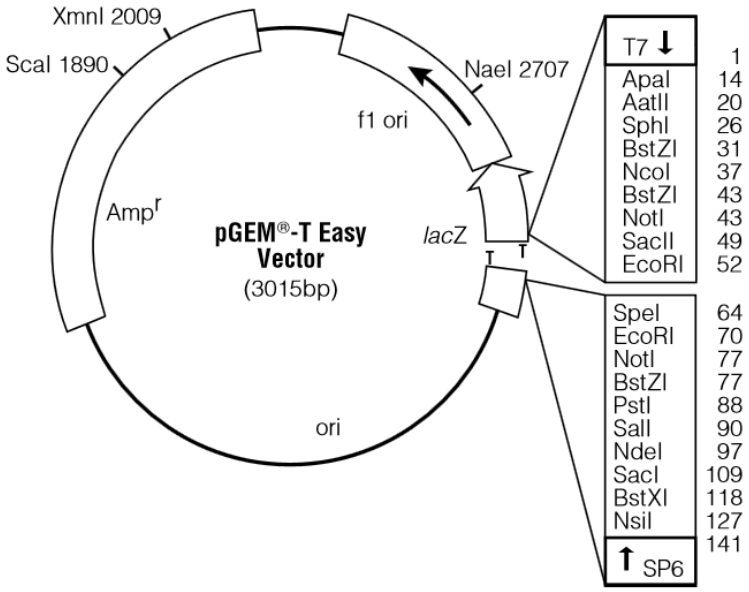


Figure 21. pGEM-T easy vector.

1.3.2. pGEX-4T-2

The pGEX-4T-2 vector was used for the heterologous expression of VvSDHs in *Escherichia coli* BL21 (DE3). The inserted sequence is under the control of IPTG inducible tac promoter and fused to the GST sequence. The produced recombinant protein is fused to the GST protein (26 kDa) at its N-terminal.

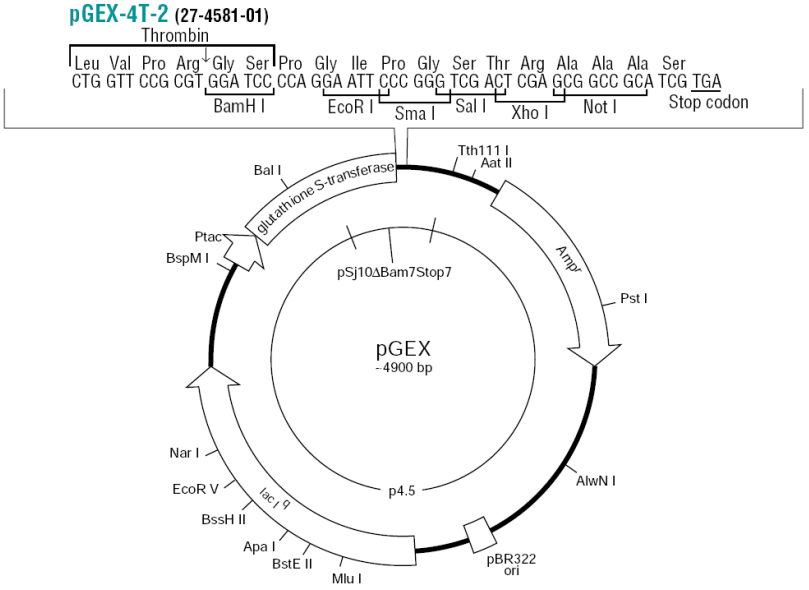


Figure 22. pGEX vector.

1.3.3. pENTR™/D-TOPO®

The vector pENTR™/D-TOPO® (Invitrogen) allowed directional cloning of VvSDHs cDNA. Bacteria transformed with pENTR™/D-TOPO® were grown in the presence of kanamycin (50 µg.mL^{-1}).

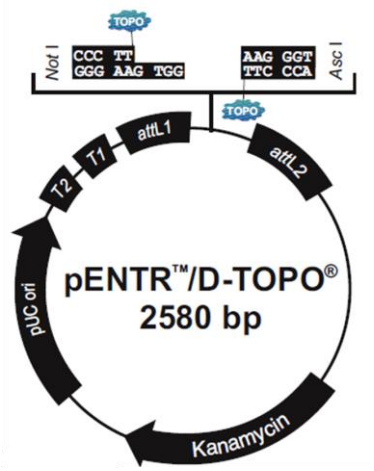


Figure 23. pENTR/D-TOPO vector.

1.3.4. pCR™8/GW/TOPO®

The vector pCR™8/GW/TOPO® (Invitrogen) was used as the entry vector to clone the cDNA of VvGATs before passing to the destination vector pEAQ-HT-DEST1. This vector was chosen because it does not confer kanamycin resistance as the destination vector pEAQ-HT-DEST1, and therefore allowed the screening of clones after the LR reaction. Bacteria transformed with pCR™8/GW/TOPO® were grown in the presence of spectinomycin (50 µg.mL^{-1}).

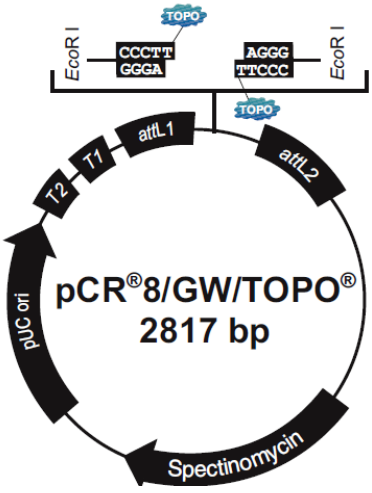


Figure 24. pCR8/GW/TOPO vector.

1.3.5. Expression vector pH2GW7

The expression vector pH2GW7 was used to clone the cDNA of VvSDHs and VvGATs. This is a binary vector of the Gateway® technology (Karimi *et al*., 2002). The sequence cloned into this vector is under the control of the 35S promoter from cauliflower mosaic virus, allowing its overexpression. The pH2GW7 vector carries the lethal gene *ccdB* between *att*R1and *att*R2 recombination sequences. Bacteria transformed with pH2GW7 vector were grown in the presence of spectinomycin (50 µg.mL^{-1}). This vector was used for stable transformation of grapevine hairy-roots and transient transformation of grapevine leaves.

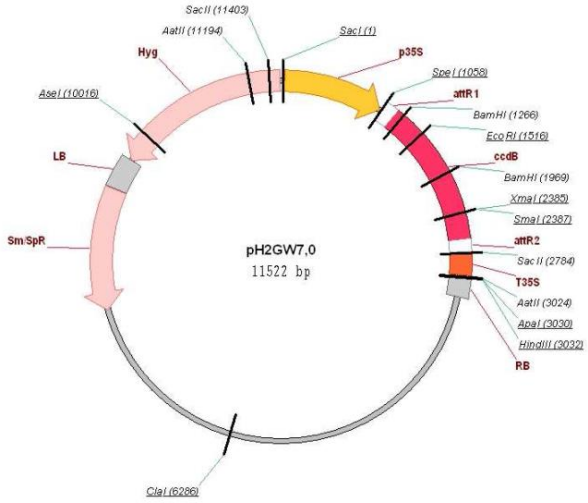


Figure 25. pH2GW7 vector.

1.3.6. Expression vectors pEAQ series

pEAQ vectors are a series of binary vectors optimized for transient expression in plants (Sainsbury *et al*., 2009). pEAQ-HT-DEST1 was used for *VvGAT1* and *-2* cloning and transient expression in tobacco leaves. pEAQ-HT-GFP was used as a control.

Table 11. List of vectors.

Vector name	**Cloning technic**	**Selection antibiotic used**	**Length (bp)**
pGEM®-T Easy	Restriction enzymes		3015
pGEX-4T-2	Restriction enzymes	Ampicillin	~4900
pENTR™/D-TOPO®	GATEWAY	Kanamycin	2580
pCR™8/GW/TOPO®	GATEWAY	Spectinomycin	2817
pH2GW7	GATEWAY	Spectinomycin	11522
pEAQ-HT-GFP	GATEWAY	Kanamycin	
pEAQ-HT-DEST1	GATEWAY	Kanamycin	11654

2. NUCLEIC ACIDS.

2.1. Extraction.

2.1.1. Plasmid DNA extraction (miniprep).

The plasmid DNA of *E.coli* was extracted with High purity plasmid Miniprep kit (Neo Biotech) according to the supplier's instructions. 1.5 mL of overnight incubated bacterial solution was centrifuged (13,200 rpm, 5 min) to recover the bacterial cells in the pellet. The pellet was resuspended in 250 µL of Resuspension Solution by pipetting. 250 µL of Lysis Solution was added and the tube was mixed by inversion. 350 µL of Neutralization Solution was added and the tube was mixed again by inversion. After centrifugation (13,200 rpm, 10 min), the supernatant was transferred to a mini-column fitted into a collection tube for further centrifugation (13,200 rpm, 1 min). After removal of the liquid passed through the membrane, 500 µL of Wash Buffer PB was loaded onto the column before centrifugation (13,200 rpm, 1 min). After removal of the liquid passed through the membrane, 500 µL of Wash Buffer W were loaded onto the column prior to centrifugation (13,200 rpm, 1 min). This washing with Wash Buffer W was repeated again. The collection tube was centrifuged for additional 1 min to remove residual ethanol. The column was finally transferred to a 1.5 mL Eppendorf tube and 30 µL of preheated to 60°C elution buffer (10 mM Tris-HCl, pH 8.5) were loaded onto the center of the column. After incubation for 2 min at room temperature, the column/collector tube was centrifuged (13,200 rpm, 2 min) and the liquid passed through the membrane was collected. The plasmid solution was stored at -20°C.
After linearization with the adequate restriction enzyme, 2 µL of plasmid solution were loaded in an agarose gel to check plasmid length and concentration.

2.1.2. RNA extraction.

The RNA extraction protocol of tobacco leaves is adapted from the RNeasy Plant Mini Kit (Qiagen, Courtaboeuf, France). 100 mg of plant tissue powder were taken up in 400 µL of lysis buffer supplemented with 1% β-mercaptoethanol. The solution was vortexed and incubated 3 min in dry-bath at 56°C before being transferred to a provided QIAshredder column, fitted in a 2 mL collection tube. After centrifugation (2 min, 13,200 rpm) to remove the fragments of plant tissue, the liquid passed through the membrane was transferred in a 1.5 mL Eppendorf tube. A half volume of ethanol was added and mixed by pipetting to the recovered liquid to precipitate nucleic acids. The mixture was loaded onto a provided RNeasy column, fitted in a 2 ml collection tube and centrifuged (15 s, 10,000 rpm) before removing the liquid passed through the column. 350 µL of washing buffer were added to rinse the RNeasy column by centrifugation (15 s, 10,000 rpm). After removing the liquid passed through the column, DNase solution (70 µL DNA digestion buffer, 10 µL DNase I) was loaded onto the center of the column and incubated for 15 min. 350 µL of buffer RW1 were again placed in the column before centrifugation (15 s, 10,000 rpm). After removing the liquid passed through the column, 500 µL of a second washing buffer were added in the column, which was centrifuged (15 s, 10,000 rpm). This step was repeated with a longer centrifugation (2 min instead of 30 s) to remove residual ethanol, then the column was transferred into a new collection tube. Another centrifugation (1 min, 13,200 rpm) was performed before transferring the column in a 1.5 mL Eppendorf tube. After addition of 35 µL of supplied water and incubation for 1 min, the tube was centrifuged (1 min, 13,200 rpm).
The quality and the quantity of purified RNAs were measured using the Nanodrop 3300 (ThermoScientific, Wilmington, USA). The solution absorbance was measured at OD_{230nm}, OD_{260nm} and OD_{280nm}. An OD_{260nm}=1 corresponds to 40 µg/mL RNA. The ratios OD_{260nm}/OD_{230nm} and OD_{260nm}/OD_{280nm} > 1.7 attest the purity of the RNA solution. Lower values indicate potential contamination with phenolic compounds or proteins.

2.2. Polymerase Chain Reaction (PCR)

2.2.1. Complementary Deoxyribonucleic Acid synthesis (cDNA).

The ImProm-II™ Reverse Transcription System kit (Promega) was used for the production of cDNA from RNA.
1 µg RNA was mixed with 1 µL Oligo(dT)15 and 10 µL Nuclease-free water in a PCR tube. The tube was placed in the heat block of the thermocycler (70°C) stopped after 5 min of heating. Once the temperature had dropped to 25°C, the tube was placed on ice during 5 minutes. The reverse transcription reaction mix was prepared in the order list with 0.8 µL Nuclease-free water, 4 µL ImProm-II™ 5X Reaction Buffer, 1.2 µL MgCl2, 1 µL dNTP Mix and 1 µL ImProm-II™ reverse transcriptase, and added to the PCR tube. The PCR tube was replaced in the thermocycler to follow the program 5 min at 25°C, 1h at 42°C, and 15 min at 70°C. Finally, 180 µL Nuclease-free water was added and the PCR tube was stored at -20°C.

2.2.2. Polymerase Chain Reaction (PCR).

The PCR allows the amplification of a DNA sequence using specifically designed primers. PCR was used both for cDNA amplification before cloning and for positive clones screening.

For cDNA cloning in vectors, PCR was monitored using the Advantage-HF2® kit (High Fidelity, Clontech, USA).
The template (cDNA, DNA sequence cloned in a vector) was amplified using 1 µL of each primer (forward and reverse, 5 µM), 2.5 µL 10X HF 2 PCR Buffer, 2.5 µL 10X HF 2 dNTP Mix, 0.5 µL 50X Advantage-HF 2 Polymerase Mix and PCR-Grade H_2O *qs* 25 µL.
The following steps were programmed on the thermocycler.
The first step consists in DNA denaturation (3 min at 95°C). The 3 following steps: 30s at 95°C, 30s at X°C for primer annealing depending on primer sequences, 1 min 30s at 68°C for Taq Polymerase elongation, were repeated 39 times in this order. After an elongation step (6 min at 68°C), the temperature was lowered to 10°C.

For colony screening, PCR was monitored using the GoTaq® DNA Polymerase kit (Promega). Single bacterial colonies grown on Petri dishes were picked with a sterile pipet tip and briefly deposited in a PCR tube containing 5 µL sterile water. In parallel, the inoculated pipet tip was finally introduced in 3 mL LB supplemented with the adequate antibiotic for overnight incubation under shaking (37°C, 220 rpm). Primers were designed for each type of plasmid used, flanking the cloning site. The orientation of the cloned sequence has been checked combining plasmid and gene primers.
5 µL 5X Green GoTaq® Reaction Buffer, 0.5 µL dNTP, 1 µL each primer (5 µM), 0.125 µL GoTaq® DNA Polymerase were added to the PCR tube. PCR was monitored with the following program: 3 min at 95°C, 30s at 95°C, 30s at X°C (depending on the primer sequence), 1min 30s at 72°C, 2 min at 72°C and 10°C until samples recovery.

2.2.3. Quantitative PCR.

Quantitative PCR (qPCR) primers were designed using Primer 3 with default parameters (Rozen & Skaletsky, 2000), to produce 120-200 bp long amplicons.

Their specifity for the selected gene has been checked by blast on the *Vitis vinifera* genome browser (http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/).
qPCR was monitored using the SYBR® Green PCR Master Mix kit (Applied Biosystems). 5 µL of template (cDNA) was mixed with 10 µL 2X SYBR® Green PCR Master Mix, 1 µM of each primer (5 µM) and 3 µL sterile water. The Real-Time (RT) PCR was monitored with 7300 Real-Time PCR System (Applied Biosystems) and analyzed with 7300 System SDS Software v 1.3.1.
To normalize the gene expression, *EF1α* was used as reference gene (Terrier *et al.*, 2005; Monteiro *et al.*, 2013). Relative gene expression was calculated from cycle threshold (Ct) according to the formula $2\ \exp-(Ct-Ct_{EF1\alpha})$.

2.2.4. Gel electrophoresis.

Electrophoresis on agarose gel was used to check the length of PCR amplified DNA fragments, released plasmid after cleavage with restriction enzymes, and linearized vectors. 1 g agarose was diluted in 100 mL of TAE (Tris acetate EDTA (Ethylene diamine tetra acetic Acid)) buffer 1X (40 mM Tris, 20mM acetic acid, 1mM EDTA, pH 8) (Biosolve). This solution was heated in a microwave until complete dissolution to obtain a homogeneous gel at 1% agarose (m/v), supplemented with Ethidium Bromide (0.5 $\mu g.mL^{-1}$, Sigma). The liquid agarose gel was poured into a specific mould equipped with combs to form wells after solidification. The solid gel was horizontally placed in an electrophoresis cell filled with 1X TAE buffer. Samples were mixed with a drop of loading buffer (if necessary) and then filled in separated wells. A well was systematically filled with 10 µL of SmartLadder (0.2 to 10 kbp, Eurogentec) to check DNA length and concentration. DNA migration was performed at 100V during a variable time (10-30 min) depending on the gel thickness and DNA length. Agarose gel was placed under a Ultra-Violet transilluminator to check DNA length and concentration.

2.2.5. DNA purification.

Gel Extraction / PCR and DNA fragment purification kit (Neo Biotech) was used to purify DNA.
For DNA extraction from agarose gel, a gel slice containing the DNA band to purify was weighted and placed in a 1.5 mL Eppendorf tube. 1µL Gel Solubilization Buffer (BD) per mg of gel was added in the tube before incubation at 60°C in water-bath until complete dissolution (about 10 min).
For DNA extraction from PCR, an equivalent volume of Solution BD was added in the PCR tube and mixed by vortexing.
The following procedure was the same for each kind of extraction.
The mixture was transferred in a Quick Gel Extraction Columns assembly and incubated 2 min at room temperature before centrifugation (12,000 rpm, 1 min). After removing the flow-through, the column was washed with 500 µL Wash Buffer (PE), previously diluted with 100% ethanol, added in the column and the assembly was centrifuged (12,000 rpm, 1 min). After removing the flow-through, the column was washed a second time in the same way. The assembly with empty column was then centrifuged (12,000 rpm, 3 min) lid open to remove residual wash buffer. The column only was transferred in a 1.5 mL Eppendorf tube and 30 µL of Elution Buffer (10 mM Tris-HCl, pH 8.5, preheated at 65°C) was added in the center of the column. After 2 min of incubation at room temperature, the assembly was centrifuged (12,000 rpm, 1 min). The flow-through containing the DNA solution was recovered and stored at -20°C.

2.3. Cloning

2.3.1 Cloning with restriction enzymes.

The primers were designed to include recognition site of restriction enzymes. BamHI recognition site (GGATCC) was added at 5'-end of the forward primer. XhoI recognition site (CTCGAG) was added at 3'-end of the reverse primer. After PCR amplification of the target cDNA with Advantage-HF2® kit, the PCR product was loaded into an agarose gel, to check its length, and purified. The purified amplicon was loaded into an agarose gel to check its length and estimate its concentration. The volume of insert to use was calculated to obtain a 3:1 molar ratio of insert:vector. 150 ng insert was mixed with 1 µL pGEM®-T Easy Vector (50 ng), 5 µL 2X Rapid Ligation Buffer, 1 µL T4 DNA Ligase, sterile water *qs* 10 µL. The reaction was mixed by pipetting, incubated for 1 h at room temperature, and then used for heat shock transformation in DH5α competent cells. The classic protocol was modified resuspending the final pellet with the rest of supernatant, 40 µL X-gal (50 $mg.mL^{-1}$) and 10 µL IPTG (100 mM), before being spread over a Petri dish of LB supplemented with ampicillin.
pGEM®-T Easy vectors containing the insert were extracted by miniprep from PCR screened positive clones.
pGEM®-T Easy vector containing the insert and the destination vector pGEX-4T-2 were previously digested with the restriction enzymes BamHI and XhoI before ligation of sticky ends with T4 DNA Ligase (Roche, Mannheim, Germany). 150 ng of digested pGEX-4T-2 vector was mixed with X ng insert released from pGEM®-T Easy vector (3:1 molar ratio), 3 µL 10X Ligation buffer, and 1U T4 DNA Ligase, sterile water *qs* 30 µL. The mixture was incubated overnight at 16°C before heat shock transformation in BL21 (DE3).

2.3.2. GATEWAY Cloning.

The forward primer was designed to include the 4 base pairs sequence CACC at the 5' end, necessary for directional cloning. After PCR amplification of the target cDNA with Advantage-HF2® kit, the PCR product was loaded into an agarose gel, to check its length, and purified. The purified amplicon was loaded into an agarose gel to check its length and estimate its concentration.

The pENTR™ Directional TOPO® Cloning kit was used to clone the insert in pENTR™/D-TOPO® vector. The volume of insert to use was calculated to obtain a 2:1 molar ratio of insert:vector, and mixed with 1 µL Salt Solution, 1µL pENTR™/D-TOPO® vector, water *qs* 6 µL in a 1.5 mL Eppendorf tube. The mix was incubated 1h at room temperature before heat shock transformation in *E.coli*.

The transfer of the insert from entry to destination vector was made using the Gateway®LR Clonase™ II Enzyme Mix (Invitrogen). 100 ng of entry vector containing the insert was mixed with 100 ng destination vector, TE buffer pH 8 *qs* 8 µL. After adding 2 µL LR Clonase™ II enzyme mix, the mixture was incubated 1 h at room temperature. 1 µL Proteinase K solution was added to the mixture and then incubated 10 min at 37°C for enzyme degradation. The mixture was used for heat shock transformation in *E.coli*.

2.3.3. Nucleotide sequencing.

Once cloned in a vector, the presence of the insert has been checked by PCR using specific primers.
Then, an aliquot of the vector solution and the insert specific primers have been sent for Sanger sequencing (GATC Biotech, Konstanz, Germany).
The obtained sequences have been blasted on the *Vitis vinifera* genome browser to check their homology.

3. TRANSFORMATION

3.1. Heat shock transformation in *E.coli*.

The kit Library Efficiency® DH5α™ Competent Cells (Invitrogen) was used for transformation in *E.coli*.
50 µL of competent cells were thawed on ice and then mixed with the ligation or LR reaction in a 1.5 mL Eppendorf tube and then incubated 30 min on ice. The tube was placed in dry-bath at 42°C during 45s and placed on ice during 5 min. 900 µL of SOC medium was added in the tube and incubated 1 h at 37°C under agitation (220 rpm). After centrifugation (8,000 rpm, 25 s), 600 µL of supernatant was removed and the pellet was resuspended with the rest of supernatant. The solution was finally spread over a Petri dish containing solid LB medium supplemented with the adequate antibiotic, and incubated overnight at 37°C.

3.2. Preparation of competent agrobacteria

One handle of agrobacteria grown on Petri dish was used to inoculate a pre-culture of 3 mL liquid LB supplemented with the adequate antibiotic incubated overnight under shaking (28°C, 180 rpm). 1 mL grown overnight pre-culture was used to inoculate a 100 mL liquid LB autoclaved Erlenmeyer supplemented with the adequate antibiotic. The culture was incubated in the same conditions till it reached OD_{600nm}= 0.5 and then was divided in two 50 mL tube for centrifugation (4,000 g, 10 min, 4°C). Each pellet was resuspended with 6 mL glycerol 10% and centrifuged in the same conditions. Each pellet was resuspended in 2 mL glycerol 10%, pooled and the solution was centrifuged in the same conditions. The final pellet was resuspended in 0.5 mL glycerol 10% and divided in 50 µL aliquots in 1.5 mL Eppendorf tubes. The bottom part of the Eppendorf tubes was immersed few seconds in liquid nitrogen before storage at -80°C.

3.3. Transformation in competent agrobacteria

An aliquot of 50 µL of competent agrobacteria were thawed on ice. Approximately 30 ng of the plasmid were added to this thawed aliquot. The resulting mixture was transferred in an electroporation cuvette placed in ice. The cuvette was placed in an electroporator to undergo discharge (1.5 kV, 25 µF, 200 Ω) during few ms. 1 mL of appropriate culture medium was added to resuspend the just transformed agrobacteria solution and transferred to a culture tube placed under agitation (180 rpm, 28°C) for 2 h. The solution was spread over a culture dish containing solid medium and appropriate antibiotic(s), and then placed at 28°C for 2-3 days.

3.4. Stable transformation of grapevine plantlets

10 µL of agrobacteria A4 (from the glycerol stock) transformed with pH2GW7 vector were spread over a solid MGL/B spectinomycin dish, and then placed 2-3 days at 28°C. One handle of agrobacterium grown on Petri dish was used to inoculate 5 mL of liquid MS/2 containing vanillin (600 µM). The amount of cells in the inoculum was measured at 600 nm with a spectrophotometer and adjusted to 0.3. *In vitro* grown grapevine seedlings had a minimum of 5 nodes to be inoculated. The caulinar apex and the last 2 emerged leaves were detached. The released limbs were pinched during 2-3 seconds with sterile long forceps previously soaked in the agrobacteria solution. The tube containing the inoculated seedlings was locked and sealed with Parafilm to conserve moisture and replaced in the greenhouse. Approximately 10 days post-inoculation, the Parafilm was removed. The time to onset of callus at the inoculated areas varies (from 1 week). The generated calli were transplanted on Petri dishes containing LGO medium (Torregrosa & Bouquet, 1997, see annexes). Up to 5 calli were grown in the same Petri dish until appearance of hairy-roots. Each appeared root was transplanted on LGO medium Petri dish and treated as an independent transformation event. The hairy-roots were maintained in culture through monthly subcultures.
After several subcultures, hairy-roots were extracted from agar medium, cleaned, weighted and frozen in liquid nitrogen. The samples were cold-milled using a 6770 Freezer/Mill® cryogenic mill (SPEX® SamplePrep, USA) into a fine powder.

3.5. Transient transformation of grapevine leaves

The protocol described in this paragraph is inspired from transient transformation of grape leaves developed at INRA in Colmar (Santos-Rosa *et al.*, 2008). 10 µL of transformed *Agrobacterium tumefaciens* C58C1 from glycerol stock were spread over a solid LB Petri dish containing tetracyclin (5 $\mu g.mL^{-1}$) and an additional antibiotic (50 $\mu g.mL^{-1}$) for vector selection: kanamycin (pEAQ-HT-GFP vector) or spectinomycin (pH2GW7 vector). The inoculated Petri dish was incubated 2-3 days at 28°C. A single colony allowed to inoculate 5 mL liquid LB containing the adequate antibiotics. After 2 days of incubation at 28°C under shaking, 1 mL of pre-culture was used to inoculate 5 mL liquid LB containing 10 mM MES (pH 5.6), 150 µM acetosyringone and appropriate antibiotics. After one additional day of incubation under the same conditions, the culture of agrobacteria was centrifuged (3,500 rpm, 15 min). The pellet was resuspended in 5 mL of infiltration buffer (10 mM MES pH 5.6, 10 mM $MgCl_2$, 2% (w/v) sucrose, 150 µM acetosyringone) and left a few hours at room temperature. The cell concentration of the agrobacteria solution was measured with a spectrophotometer at 600nm. The $OD_{600\ nm}$ was adjusted to 0.4 with infiltration buffer. The vitroplants used for leaves agroinfiltration were grown at INRA in Colmar. The experiments were performed there. Seedlings were grown in glass tubes containing 20 mL of Galzi medium (Galzi, 1964) in a growth chamber (21°C, photoperiod: 18h day / 6h night). 8 to 10 weeks old seedlings were used for this experiment. The last 2 emerged leaves of each vitroplant were detached and placed on a Petri dish containing a sterile Whatman paper disc hydrated with sterile water, upper side against the paper. The leaves of 6 seedlings were used for each genetic construct. The two detached leaves of the same seedling were placed in a flask containing the agrobacteria solution, lower face in contact with the solution, and covered with a disk of Miracloth. The flasks were placed in a vacuum bell jar and subjected to the vacuum pressure exerted by a pump (KNF Neuberger GmbH, N035.3AN.18 type) for 1 min. The bell was manually agitated during application of the vacuum to prevent the formation of air bubbles. The vacuum procedure was repeated twice. After a rinsing step in sterile water, the leaves were incubated again on hydrated Whatman paper and the Petri dish was sealed with Parafilm to retain moisture. The dishes containing the transformed leaves were kept in a growth chamber. Grapevine leaves were frozen and ground in a mortar with a pestle in the presence of liquid nitrogen, 4 to 5 days after

inoculation.

3.6. Transient transformation of tobacco leaves

Ten µL of glycerol stock of transformed *A. tumefaciens* GV3101 was spread on a solid LB Petri dish containing kanamycin (50 µg.mL^{-1} for pEAQ vectors selection) and placed for 2-3 days at 28°C. One handle of agrobacteria was used to inoculate 5 mL liquid LB containing kanamycin (50 µg.mL^{-1}) incubated at 28°C under shaking (180 rpm). After overnight incubation, the pre-culture was centrifuged (10 min, 3,200 rpm, room temperature). The pellet was resuspended in 1 mL of infiltration buffer (MES-KOH pH 5.6 10 mM, 10 mM $MgCl_2$, 200 µM acetosyringone). The $OD_{600\ nm}$ was measured with a spectrophotometer (Cary 100 Bio UV-Visible, Varian) from an agrobacteria solution diluted 20 times. The final OD was adjusted to 0.5 with the infiltration buffer. The agrobacteria solution was incubated at room temperature for a few hours before infiltration at the underside of tobacco leaves using a 10 mL plastic needle. 4-6 weeks old tobacco plants (*Nicotiana benthamiana*) were infiltrated, favouring not damaged leaves. 3-4 leaves per plant were used and the infiltrated areas were marked by a permanent marker. 4-7 days after inoculation, the infiltrated areas were harvested, weighted and frozen in liquid nitrogen.
In some cases, a substrate mix was infiltrated in the same way, 3 days after agroinfiltration. Certain leaves were detached just after substrates infiltration and their petiole was dipped overnight in an Eppendorf tube containing the substrates mix. For these experiments, the leaves were harvested the following day.
The samples were cold-milled using a 6770 Freezer/Mill® cryogenic mill (SPEX® SamplePrep, USA) into a fine powder.

4. PROTEINS

4.1. Heterologous expression in *E.coli*

Liquid LB medium containing ampicillin (100 µg.mL^{-1}) was inoculated with *E.coli* BL21 cells containing *VvSDH* cDNA fused with GST sequence in the pGEX-4T2 vector. The bacterial solution was incubated overnight at 37°C under shaking (220 rpm). 3mL of the pre-culture were used to inoculate an erlenmeyer containing 120mL of YTA 2X medium with ampicillin (100 µg.mL^{-1}). The Erlenmeyer was incubated at 37°C under shaking and bacteria growth was monitored using absorbance at 600nm with a spectrophotometer.

Induction of the synthesis of the recombinant proteins: Once OD_{600} was between 0.8 and 1, 0.1 mM final IPTG (isopropyl-1-thio-β-D-galactopyranoside) was added to trigger the induction of GST-SDH fusion proteins. The Erlenmeyer was placed at room temperature under slow shaking during about 20 hours.

Preparation of bacterial cell lysate: After this induction step, culture medium was centrifuged (8000g, 10 min, 4°C). The pellet was resuspended in cold PBS, 0.1% (v/v) β-mercaptoethanol, 1% (v/v) Triton X-100, and lyzosyme (final concentration of 1 mg.mL^{-1}). The solution was sonicated in cold conditions to break cell membranes and then centrifuged (12,000 g, 5 min,10°C) to recover the supernatant (S). An aliquot of the pellet (P) and the supernatant was conserved for further analysis by SDS-PAGE (Sodium dodecyl sulfate- polyacrylamide gel electrophoresis).

Enzyme purification: After adding Glutathione-Sepharose™4B (GE Healthcare, Chalfont St Giles, UK), the supernatant was set to slow shake on rotary axis for 1 h at room temperature. After centrifugation (1,500 g, 5 min, 10°C), the supernatant, containing non linked (NL) molecules, was conserved for further analysis by SDS-PAGE gel. The pellet was resuspended in PBS with 0.1% (v/v) *β*-mercaptoethanol to wash the Glutathione-Sepharose 4B beads. After centrifugation (1,500 g, 5 min, 10°C), the supernatant was designated Wash 1 (W1) and conserved. Two supplementary were performed in the same conditions and the supernatant was systematically recovered (W2 and W3). A thrombin solution was added to the resin and liberated proteins. After slow shake on rotary axis for 3 h at room temperature, the solution was centrifugated (1,500 g, 5 min, 10°C). The supernatant was collected and conserved as Eluate 1 (E1). The beads pellet was resuspended with 300 µL PBS with 0.1% (v/v) *β*-mercaptoethanol. Two additional centrifugations allowed recovery of E2 and E3.

Proteins were assayed according Bradford method (Bradford, 1976) using a bovine serum albumin reference curve. Purity and molecular weight of the recombinant proteins were checked by running an aliquot on SDS-Polyacrylamide gel (14%) by electrophoresis. Recombinant proteins were diluted in glycerol 16 % (v/v), aliquoted and stored at -20°C.

4.2. Proteins extraction from tobacco leaves

Different buffers and desalting/buffer exchange columns were used.
The column equilibration buffer was the same as the buffer used for further enzymatic reaction.
For each method, tobacco leaves powder was mixed with extraction buffer in a 1.5 mL Eppendorf tube at a 1:3 (w:v) ratio. Once mixed by vortexing, the tube was placed on a rotary axis at 4°C during 1h before centrifugation (30 min, 15,000 g, 4°C).

Method 1:
The extraction buffer was composed of MES (2-(*N*-morpholino)ethane sulfonic acid) pH 6.5 (adjusted with KOH) 50 mM, NaCl 150 mM, PVPP (Polyvinylpolypyrrolidone) 1% (w:v), Triton X100 1% (v:v).
A PD Mini-trap G25 column (GE Healthcare) was equilibrated with MES-KOH buffer 50 mM (equilibration buffer), following the manufacturer recommendations. The supernatant was recovered and loaded onto the column. 200 µL equilibration buffer were added several times in the column for elution. Eluate fractions (6-7 x 200 µL) were recovered in 0.5 mL tubes at the column output and stored on ice. Total proteins concentration was assayed with Bradford method for each fraction. The most concentrated fractions were used for enzymatic assays.

Method 2:
The extraction buffer was composed of MES-KOH pH 5.6 50 mM, NaCl 150 mM, PVPP 1% (w:v), Triton X100 0.1% (v:v). Enzyme extract was desalted with PD Mini-trap G25 column.

Method 3:
The extraction buffer was composed of MES-KOH pH 6.5 50 mM, NaCl 150 mM, PVPP 1% (w:v) (Mugford *et al*., 2009).
900 µL supernatant and MES-KOH (pH 6.5 25 mM) *qs* 2 mL were loaded onto an Amicon® Ultra-4 column (Molecular weight cut-off: 10 kDa, Merck Millipore). The column was centrifuged 3 times (40 min, 4,000 g, 4°C). After each centrifugation, MES-KOH (pH 6.5 25 mM) *qs* 2 mL were added onto the column. The final volume (30-50 µL) retained in the column was used for enzymatic assay.

4.3. Protein assay according to the Bradford method.

Proteins produced in the different experiments were assayed according to the Braford method.
Bradford reagent was prepared as follows: 100 mg Brilliant Blue G (Sigma-Aldrich) were diluted in

50 mL ethanol 95% under shaking in the dark. 100 mL H_3PO_4 were added to the solution, completed with milli-Q water *qs* 1L. Bradford reagent was stored in a bottle covered by aluminium foil, at 4°C.

A standard range was prepared before each assay from a solution of Bovine Serum Albumin (BSA) at 1 $g.L^{-1}$. One to 20 µL of this solution were poured into individual spectrophotometer cuvettes. 1 ml of Bradford reagent was added to each cuvette every 15s. The absorbance was measured using a spectrophotometer (Cary 100 Bio UV-Visible Varian) at 595nm, respecting a period of 15 s between each measurement. A standard curve was modeled using Excel from the graph OD_{595nm} = f (protein mass). Protein extract were assayed using the same approach. The obtained OD_{595nm} was plotted on the graph to determine the protein concentration in the sample.

4.4. SDS-PAGE

Protein separation was performed by SDS-Polyacrylamide gel electrophoresis (SDS-PAGE).
Protein samples were diluted twice with a loading buffer (Tris-HCl 0.125 mM, pH 6.8, SDS 4% (w/v), glycerol 20% (v/v), β-mercaptoethanol 10% (v/v), bromophenol blue 0,1% (w/v), and denaturated by heating at 70°C during 30 min.
The acrylamide gel was constituted of two superimposed gels (stacking and running gels) poured vertically in the same cassette. The stacking gel was composed of acrylamide (4 %), stacking buffer (Tris HCl 0.5 M, pH 6.8), SDS 10% (w/v), ammonium persulfate 10% (w/v) and tetramethylenediamine (TEMED). The running gel was composed of acrylamide (12%), running buffer (Tris HCl 1.5 M, pH 8.8), SDS 10% (w/v) and ammonium persulfate 10% (w/v).
The cassette was placed in an electrophoresis tank to apply a vertical electric field across the gel.
Migration was monitored in presence of migration buffer (Tris HCl 25 mM pH 8.3, glycine 192 mM, SDS 0.1% (w/v). It started in the stacking gel (20 mA, 30 min) and continued in the running gel (30 mA, 1h30).
The gel was incubated overnight in a staining solution composed of Coomassie Blue R250 0.1% (w/v), ethanol 40% (v/v), acetic acid 10% (v/v). The gel was then discoloured with acetic acid: ethanol: water (10/40/50) and finally wrapped in cellophane in presence of few water drops and stored at 4°C.

5. ENZYMATIC ASSAYS

5.1. Shikimate dehydrogenases

To determine kinetic parameters of VvSDHs, the reactions were performed at 30°C, in a final volume of 200 mL containing 100mM Bis-Tris Propane HCl buffer, and started from addition of 100 ng of enzyme in the reaction mixture.
Michaelis constant for SA ($K_{m\,(SA)}$) was determined at pH 7 and pH 9, with reaction mixture containing 2 mM $NADP^+$ and variable concentration of SA. $K_{m\,(NADP+)}$ was determined at the same pH with 4 mM SA and variable concentration of $NADP^+$. $K_{m\ (3\text{-}DHS)}$ was determined at pH 7 for VvSDH5 and VvSDH8 using NADPH as cofactor (0.4 mM).
Effect of assay pH was monitored with 10 mM of substrate and 2 mM of cofactor from 100 ng of recombinant proteins in the same buffer at pH 6.5, 7, 7.5, 8, 8.5 and 9.
Kinetic parameters were determined following NADPH variation at 340 nm by spectrophotometry (see Spectrophotometry). For 3-DHS and GA assay, 1 µg of SDH was added to a reaction mixture containing 100mM Bis-Tris Propane HCl buffer pH 9, 4 mM SA or 3-DHS, 2 mM $NADP^+$ and 8 mM ascorbic acid in a final volume of 80 µL. After 2 hours of incubation at 30°C, the reaction was stopped with 20 µL HCl 11.7 N and analysed by HPLC.

5.2. Glucosyltransferase

A grapevine glucosytransferase (VvGT2) earlier characterized in the lab (Khater *et al*., 2012) was used for the production of hydroxycinnamoyl-glucoses, potential substrates for VvGATs. VvGT2 was produced by heterologous expression in *E.coli*.
The reactional mixture contained 1 mM HCA (caffeic, ferulic or *p*-coumaric acid), 2.5 mM UDP-glucose, 100 mM MES-KOH pH 6.5 buffer, 4 mM ascorbic acid and 12 µg.mL^{-1} GT2 in 800 µL of final volume . The mixture was incubated in dry-bath at 30°C during 1h30, stopped with 200 µL acetone:water:TFA (70:30:0.05, v/v/v) and stored at -20°C until purification.

5.3. Glucose acyltransferase

The first assay was monitored in a 1.5 mL Eppendorf tube containing 1 mM β-G, 0.4 mM EC, 4 mM ascorbic acid in 50 mM potassium phosphate buffer pH 6, and about 300-400 µg desalted proteins prepared according to method 1, in a final volume of 1 mL. The mixture was incubated in dry-bath at 30°C during 1 h, stopped with 200 µL acetone:water:TFA (70:30:0.05, v/v/v) and evaporated with Genevac (HPLC fraction mode, 35°C). The pellet was resuspended with 70 µL methanol:water:TFA (20:80:0.05, v/v/v) before centrifugation (15 min, 13,200 rpm, 4°C). The supernatant was recovered for HPLC analysis.
For the other assays, the reaction mixture was not evaporated with Genevac.

The 2nd assay was performed in 100 mM MES-KOH buffer pH 5 in the same conditions as the first assay. Protein extract was prepared according to method 2.

The 3rd assay was performed in 25 mM MES-KOH buffer pH 6.5 with 1 mM β-G, 0.4 mM EC, 4 mM ascorbic acid with ~20 µg desalted proteins. Protein extract was prepared according to method 3.

6. PREPARATION OF PLANT METABOLITES

6.1. Grapevine hairy-roots and berry tissues.

Approximately 100 mg of frozen powder of grapevine hairy-root or berry tissues were weighed and transferred into a Precellys® tube (7 mL) containing beads. 500 µL of pure methanol were added directly to dilute the powder and mixed by vortexing. 3.5 mL of acetone:water:TFA (70:30:0.05, v/v/v) were then added to the extract. The sample was ground with a grinder-homogenizer (Precellys 24®) following a cycle of 3 sequences of 40s at 5,000 rpm spaced out by 40s. This cycle was repeated 3 times. 3.5 mL of homogenate were transferred to a 15 mL tube and centrifuged (5 min, 4,500 rpm, 4°C). Two fractions of 1 mL of supernatant were dried (Genevac, SP Scientific, HPLC fraction mode, 35°C) and then stored at -80°C. One of the 2 fractions was used for a direct injection onto the HPLC system (see HPLC analysis and preparation), the other one underwent a phloroglucinolysis step before injection. The phloroglucinolysis reagent was prepared with 0.25 g of phloroglucinol and 0.05 g of ascorbic acid with 5 mL of acidified methanol (final HCl concentration: 0.2 M).

The pellet used for direct injection was resuspended in 500 µL of methanol:water:formic acid (50:50:1, v/v/v) and placed in an ultrasonic bath. After 30 min, the sample was centrifuged (15 min, 15,000 rpm, 4°C). The supernatant was recovered for analysis.

The pellet used for phloroglucinolysis was resuspended in 700 µL of phloroglucinolysis reagent using a sterile needle. After vortexing, the sample was placed in a water bath at 50°C for PA depolymerisation. After 20 min of incubation, the tube was cooled in ice and 700 µL of ammonium formate buffer 200 mM were added to stop the reaction.
After centrifugation (15 min, 15,000 rpm, 4°C), the supernatant was recovered and filtered before analysis.

6.2. Grapevine leaves

About 20 mg of powder were diluted in 400 µL of acetone:water:TFA (70:30:0.05, v/v/v) buffer. The sample was placed on a rotary axis for 1 h at room temperature and then centrifuged (15 min, 14,000 rpm, 4°C). 2 supernatant aliquots of 150 µL were transferred in 1.5 mL Eppendorf tubes to be evaporated in Genevac (HPLC fraction mode, 35°C). The obtained pellet from the first aliquot was taken up with 75 µL of methanol:water:TFA (80:20:0.05, v/v/v) buffer and placed in an ultrasonic bath (sonication) for 15 min. 75 µL of the same buffer were added and a second sonication (15 min) was performed. The obtained pellet from the second aliquot was taken up in 75µL phloroglucinolysis reagent and sonicated in ultrasonic bath for 15 min. After 20 min of phloroglucinolysis reaction in a water bath at 50°C, 75 µL of sodium acetate 200 mM were added to stop the reaction. After centrifugation (15 min, 15,000 rpm, 4°C), the supernatant was recovered for analysis.

6.3.Tobacco leaves

Approximately 50 mg of frozen tobacco leaf powder were diluted in 400 µL of methanol containing 2% formic acid. After shaking for 1 hour on a rotary axis at room temperature in the dark, the sample was centrifuged (15 min, 15,000 rpm, 4°C). The supernatant was recovered for analysis.

7. METABOLITE ANALYSIS

7.1. Spectrophotometry

VvSDH catalysed reactions from the pairs of molecules SA / $NADP^+$ or 3-DHS / NADPH were followed by the NADPH concentration variation with spectrophotometer at 340 nm (Cary 100 Bio UV-Visible, Varian). Absorbance at 340nm was recorded for 2 minutes after adding the enzyme. The activity of the enzyme was modelled by Absorbance at 340nm = f (time) using Excel software (Microsoft) to determine the slope at the origin ($\Delta A/\Delta t$).

For the determination of kinetic parameters, V_i was measured at different initial concentrations (Ci) of substrate or cofactor. Initial velocity (V_i) was calculated with the NADPH molar extinction coefficient at 340 nm (ε=6.18 x 10^3 $mol.L^{-1}$).
$Vi = (\Delta A/\Delta t) / \varepsilon_{NADPH}$.

The graph V_i = f(Ci) was fitted with Michaelis-Menten hyperbola to yield K_m and V_{max} values (Hyper32 program, Version 1.0.0, 2003). V_{max} was recalculated taking in account the reaction volume (Vr=200 µL) and the quantity of enzymes introduced in the reactional volume.

The enzyme concentration $[E_0]$ ($mol.L^{-1}$) was calculated from the molar mass of the enzyme.
The turnover number (k_{cat}) was calculated as follows:
$k_{cat} = V_{max} / [E_0]$.

The catalytic efficiency was calculated as k_{cat}/K_m ratio.

7.2. HPLC analysis and preparation

7.2.1. VvSDH enzymatic reactions

5 µL of the reactional mixture was injected on an Agilent 1100 LC system (Agilent Technologies, Waldbronn, Germany) equipped with a diode-array detector (DAD). Compounds were separated using an Atlantis dc18 (250 x 2.1, 5 µm) analytical column (Waters, Milford, MA) protected by an Atlantis dc18 (10 x 2.1, 5 µm) precolumn (Waters, Milford, MA).

For VvSDH enzymatic reactions, the mobile phase consisted of water containing 0.1% formic acid (eluent A) and methanol containing 0.5 % formic acid (eluent B).The elution program had a 0.25 mL/min flow rate and started with 0% B. The linear gradients were from 0 to 40% B (0.05-11 min), from 40 to 50% B (11-15 min) and from 50 to 100% B (15-25 min), followed by washing and reconditioning of the column. The detection of the molecules has been monitored in UV at 240 and 280 nm, depending on the maximal absorbance reported in the litterature.
Pure commercial solutions (Sigma-Aldrich) of SA, 3-DHS, $NADP^+$, GA and PCA were used as standards.

For VvGT and VvGAT enzymatic reactions, the mobile phase consisted of water containing 2% formic acid (eluent A) and acetonitrile containing 18% water and 2% formic acid (solvent B). The elution program had a 0.25 mL/min flow rate and started with 0% B. The linear gradients were from 0 to 10% B (5-35 min), from 10 to 40% (35-65 min) and from 40 to 100 % (65-70 min), followed by washing and reconditioning the column. The detection of the molecules has been monitored in UV at 280 nm for flavan-3-ols and 320 nm for HCAs.

7.2.2. Preparation of GAT substrates

VvGT2 reaction mixture was injected on the semi-preparative HPLC LC 940 Varian, equipped with a UV detector at two wavelength channels (280 and 320 nm), on a (100 x 21.4 mm i.d.) Microsorb C18 column (3µm).
The mobile phase was composed of solvent A (water:acetic acid, 99:1, v/v) and solvent B (acetonitrile:water:formic acid, 80:19:1, v/v/v) at a flow rate of 10 $mL.min^{-1}$.
Gradient was as follows:

Time (min.)	%A	%B
0	90	10
10	80	20
11	20	80
12	20	80
13	90	10
15	90	10

The fraction corresponding to the glucose ester of HCA (see annexes, Production of putative substrates for VvGATs) was collected, concentrated under vacuum and lyophilized (Edwards Lyophilizer). The lyophilisate was stored in a glass vial at -20°C until use.
β-glucogallin was a kind gift of Dr Gross (retired, formerly University of Ulm, Germany).

7.2.3. Grapevine and tobacco leaves metabolites

PA units, i.e. terminal units and phloroglucinol adducts formed from extension units obtained after phloroglucinolysis were analyzed on a reversed solid phase column Atlantis DC18 (Agilent Technologies, Waldbronn, Germany) (5 µm packing, 250 mm × 4.6 mm id) protected by a pre-column of the same material (20 × 4.6 mm id) (Waters, Milford, MA) and a Phenomenex C18 cartridge SecurityGuard^{TM} (Torrance, USA) (4.0 mm x 3.0 mm id). The mobile phase was a linear gradient, composed of solvent A (water:formic acid, 98:2, v/v) and solvent B (water:acetonitrile:formic acid, 80:18:2, v/v/v) at a flow rate of 0.25 mL.min^{-1} at 30°C. Gradient, expressed as a proportion of solvent B, was as follows: isocratic for 5 minutes at 0%; 5-35 min, 0-10%; 35-65 min, 10-20%; 65-70 min, 20-100%;70-75,100-0%.

PA units were analyzed using 2 detectors in series.
The first system is a diode array detector (DAD) used at 280 and 320 nm for quantification of extension units substituted with phloroglucinol (Agilent Technologies, Waldbronn, Germany). The second sensor was a Shimadzu (Kyoto, Japan) spectrofluorometer used to obtain better precision in the quantification of terminal units. Excitation and emission wavelengths were 275 nm and 322 nm, respectively. Concentrations were calculated using the peak areas from external calibrated standards, either commercial ((+)-catechin, (-)-epicatechin, (-)-epigallocatechin, (-)-epicatechin-3-*O*-gallate), or purified in the laboratory (phloroglucinol derivatives).

Tobacco leaves metabolites were quantified wih the same system using DAD at 280nm. The mobile phase was a linear gradient, composed of solvent A (water:formic acid, 95:5, v/v) and solvent B (acetonitrile:water:formic acid, 80:15:5, v/v/v) at a flow rate of 0.25 mL.min^{-1} at 38°C, and started with 0% B. Gradient, expressed as a proportion of solvent B, was as follows: isocratic for 3 min at 0%; 2-5 min, 0-2%; isocratic 5-12 min, 2%; 12-15 min, 2-3%; 15-25 min 3-8%; 25-40 min, 8-20%; 40-45 min, 20-25%, isocratic 45-55 min, 25%; 55-70 min, 25-65%; 70-75 min, 65-0%.

7.3. UPLC-DAD-MS analysis

Identification of certain molecules was operated on a Ultra Performance Liquid Chromatography-Diode Array Detector-Mass spectrometry (UPLC-DAD-MS) system. Separations were performed using a Waters Acquity UPLC-DAD system (Milford, MA), on a (15 × 1 mm i.d.) Acquity BEH C18 column (Waters, Milford, MA; 1.7 µm), operated at 35°C. Mobile phase consisted of water/formic acid (99/1, v/v) (eluent A) and methanol/formic acid (99/1, v/v) (eluent B). Flow rate was 0.08 mL.min^{-1}.
The elution program for VvSDH enzymatic reactions was as follows: isocratic for 4 min with 1% B, 1-20% B (4-10 min), 20-98% B (10-12 min), isocratic with 98% B (12-17 min).
The elution program for identification of tobacco leaves metabolites was as follows: isocratic for 1 min with 2% B, 2-15% B (1-6.5 min), isocratic with 15% B (6.5-9 min), 15-30% B (9-12 min), isocratic with 30% B (12-14 min), 30-75% B (14-27 min), 75-95% B (27-32 min).

ESI-MS/MS analyses were performed with a Bruker Daltonics Amazon (Bremen, Germany) mass spectrometer equipped with an electrospray source and an ion trap mass analyser. The spectrometer was operated in the positive and negative ion modes (end plate off set: -500V; temperature, 200°C; nebulizer gas: 10 psi and dry gas, 5 L.min^{-1}; capillary voltage, 2.5 kV in positive mode, 4.5 kV in negative mode). Collision energy for fragmentation used for MS2 experiments were set at 1.

7.4. UHPLC-QqQ-MS analysis

Hairy-roots and grape berry metabolites were analyzed with an UHPLC-QqQ-MS method in the multiple reaction monitoring (MRM) mode.
This method, published in Lambert *et al*. (2015), has been used for higher sensitivity in quantification of GA, β-G after direct injection. Depolymerized PAs subunits were quantified after depolymerisation by phloroglucinolysis.

UHPLC-QqQ-MS analyses were carried out using an Acquity UPLC system (Waters) hyphenated to a triple quadrupole mass spectrometer (Waters) with electrospray ionization (ESI) operating in switching positive and negative modes. The UHPLC system included a binary pump, a cooled autosampler maintained at 7°C and equipped with a 5 µL sample loop, a 100 µL syringe and a 30 µL needle, and a DAD. MassLynx software was used to control the instruments and to acquire the data, and TargetLynx software was used to process the data. The column used for chromatographic separation was a reversed-phase Acquity HSS T3 1.8 µm 1.0 × 100 mm, (Waters) protected by a 0.2 µm in-line filter and maintained at 40 °C. The mobile phase consisted of 1% (v/v) formic acid in Milli-Q water (solvent A) and 1% (v/v) formic acid in methanol (solvent B). The source and desolvation temperatures were respectively set at 120°C and 450°C. Nitrogen was used as desolvation (500 $L.h^{-1}$) and cone (50 $L.h^{-1}$) gas. Argon was used as collision gas at a flow rate of 0.16 $mL.min^{-1}$. Capillary voltage was set at 3.5 kV in positive mode and 2.8 kV in negative mode. The MRM transitions are described in Table 12.
Samples were injected into the column by using the Partial Loop with Needle Overfill injection mode, with an injection volume of 1 µL. The following gradient elution program, at a flow rate of 0.17 $mL.min^{-1}$, was performed: isocratic 1% B from 0.0 to 2.0 min, linear 1%–5% B from 2.0 to 2.1 min, linear 5%–10% B from 2.1 to 8.0 min, linear 10%–28% B from 8.0 to 12.0 min, isocratic 28% B from 12.0 to 18.0 min, linear 28%–45% B from 18.0 to 22.0 min, linear 45%–99% B from 22.0 to 23.5 min, isocratic 99% B from 23.5 to 26.5 min. At the end of this sequence, the column was brought back to initial conditions with linear 99%–1% B from 26.5 to 27.0 min, then re-equilibrated with isocratic 1% B from 27.0 to 30.0 min. Two detectors were used: DAD scanning from 210 to 600 nm (resolution 1.2 nm) and the QqQ detector.

Table 12. List of quantified compounds and MRM parameters (ion mode, precursor and product ions m/z, retention times).

	Ion Mode	**m/z precursor**	**m/z quantifier**	**m/z qualifier**	**retention time (min)**
Direct injection					
Gallic acid	-	168.9	125.0	79.0	1.80
β-glucogallin	-	331.0	169.0	151.0	1.78
Phloroglucinolysis					
Gallocatechin	+	307.1	139.0	151.0	3.40
Epigallocatechin	+	307.1	139.0	151.0	4.93
(Epi)gallocatechin phloro	+	431.2	127.0	305.1	2.08
Catechin	+	291.1	139.0	123.1	4.96
Catechin phloro	+	415.1	127.1	289.1	3.56
Epicatechin	+	291.1	139.0	123.1	5.48
Epicatechin phloro	+	415.1	127.1	289.1	4.06
Epicatechin gallate	+	443.1	123.1	273.1	5.78
Epicatechin gallate phloro	+	567.2	153.1	247.2	5.01

Abbreviations: Phloro: phloroglucinol adduct.

8. STATISTICAL ANALYSIS

The significance of the results from HPLC analysis has been statistically assessed with a Student's t-test using two-sided alternative with the R software (version 3.12. R Development Core Team. 2009).

CHAPTER 3: SHIKIMATE DEHYDROGENASES

1. Introduction

The biosynthesis of GA has been debated in the scientific literature, notably to determine which precursor(s) can be used to produce this hydroxybenzoic acid (Dewick & Haslam, 1969). It was thought that it could be produced from metabolites of the phenylalanine or SA pathways (Fig. 1, manuscript #2). PCA hydroxylation was also proposed (El-Basyouni *et al*., 1964, Kato *et al*., 1968). It was postulated that in *Rhus succedanea*, GA could be synthesized from SA pathway in young leaves but from phenylalanine in older leaves (Ishikura *et al*., 1984). The fact that glyphosate triggers the accumulation of some HBAs (GA, PCA and 4-hydroxybenzoic acid) in several plants indicates that those metabolites could be synthesized from SA (Lydon & Duke, 1988). More recent retrobiosynthetic studies revealed that GA is mainly (more than 90%) synthesized by 3-DHS dehydrogenation (Werner *et al*., 1997, 2004), in agreement with a report from the 1960s (Dewick & Haslam, 1969). Muir and co-workers (2011) have cloned a walnut DQD/SDH able to produce GA from SA and 3-DHS in presence of $NADP^+$. These reactions were slow (overnight incubation) compared to the catalytic properties of this enzyme using SA and NADPH. Since, GA biosynthesis in plants has not been further investigated.

DQD/SDH is a bifunctional enzyme recognized to ensure the carbon flux in the SA pathway (Moore, 2004). As described in chapter 1, the DQD/SDH is able to catalyse the 3^{rd} (3-dehydroquinate dehydration) and 4^{th}steps (NADPH-dependent reduction of 3-DHS) of the SA pathway.

However, additional catalytic properties have also been described. Its SDH domain can be involved in quinic acid biosynthesis from 3-DHQ (Guo *et al*., 2014) or in GA biosynthesis from 3-DHS (Muir *et al*., 2011).

Hence, this enzyme could also produce, upstream from the flavonoid biosynthetic pathway, GA necessary to flavan-3-ol galloylation (Akagi *et al*., 2009).

An interspecific analysis of the SA pathway genes was conducted in algae, mosses, monocots and dicots whose genome has been sequenced (Tohge *et al*., 2013). This study highlightedgene duplication events which reflect the adaptation of photosynthetic organisms to terrestrial life. Indeed, the algal species have only one gene for each step of the pathway. On the contrary, terrestrial plants often possess several isoforms of some genes of the SA pathway, which is particularly true for the DQD/SDH. Tandem duplication events have been identified concerning the DQD/SDH genes in several dicotyledonous plants: castor oil plant (*Ricinus communis*), cassava (*Manihot esculenta*), cacao tree (*Theobroma cacao*), strawberry (*Fragaria vesca*) and grapevine (*Vitis vinifera*).

The grapevine genome has 4 genes encoding DQD/SDH. A first gene (designated *VvSDH5*) is located on chromosome 5. The other 3 genes (designated *VvSDH7*, *-8* and *-9*) are located side by side as a cluster on chromosome 14.

A quantitative genetic analysis (Huang *et al*., 2012) and a transcriptomic analysis from grapevine hairy- roots transformed with MybPA1 or MybPA2 (Terrier *et al*., 2009) led us to focus on *VvSDH*s. Indeed, *VvSDH9* is part of a list of genes potentially involved in the biosynthesis of PAs (Carrier *et al*., 2013).

We have tested the involvement of those 4 genes in GA metabolism by *in vitro* and *in planta* experiments.

This chapter describes the functional characterization of those 4 VvSDHs. Manuscript #1 corresponds to an article in preparation.
Other research results not included in manuscript #1 are also presented in this chapter.

2. Manuscript #1

GRAPEVINE SHIKIMATE DEHYDROGENASES INVOLVEMENT IN GALLIC ACID BIOSYNTHESIS

Thibaut Bontpart, Thérèse Marlin, Sandrine Vialet, Jean-Luc Guiraud, Lucie Pinasseau, Emmanuelle Meudec, Véronique Cheynier, Nancy Terrier.

INRA, UMR 1083 Sciences pour l'œnologie, 2 place Pierre Viala, F-34060 Montpellier cedex 1, France

In vascular plants, diverse biological roles are attributed to proanthocyanidins (PAs), also called condensed tannins. Notably, those flavonoids are involved in response against biotic and abiotic stresses (Dixon *et al*., 2005). Moreover, their accumulation and oxidation in the seed coat influence seed coloration and germination in *Arabidopsis thaliana* (Debeaujon *et al*., 2000). PAs are found in numerous fruits and foods (wine, tea, chocolate). They constitute the major flavonoids in grape berry and are further found in wine, influencing its astringency, bitterness and color stability (Arnold *et al*., 1980). Numerous epidemiological studies have shown the beneficial effects of PAs on human health. Notably, those polyphenols may play a role in the prevention of cardiovascular diseases (Kruger *et al*., 2014).

In *Vitis vinifera*, PAs are mainly type B polymers of flavan-3-ols. Indeed, C4-C8 link is the major link between the flavan-3-ol units. Flavan-3-ol monomers and PAs are found in the hypodermal layers of the skin. They are also found in the epidermis, the outer and inner integuments of the seeds (Cadot *et al*., 2006). The main extension units differ according to the tissue: (+)-catechin and (-)-epicatechin in seeds, (-)-epicatechin and (-)-epigallocatechin in skin (Souquet *et al*., 1996, Bogs *et al*., 2005). (-)-epicatechin can be acylated with gallic acid (GA), a trihydroxybenzoic acid, to form (-)-epicatechin 3-*O*-gallate. The percentage of (-)-epicatechin 3-*O*-gallate in a PA chain defines the galloylation rate (%G). %G is higher in seeds than in skin (Cheynier, 2005, Mané *et al*., 2007, Verriès *et al*., 2008). Galloylation increases antioxidant (Da Silva *et al*., 2003), antiproliferative (Lizarraga *et al*., 2007), astringency (Vidal *et al*., 2002) and colloidal (Fabre *et al*., 2010) properties of PAs.

Grape berry growth is characterized by a double sigmoid curve (Coombe & Hale, 1973). The véraison marks the beginning of the ripening stage. PA biosynthesis starts before flowering and their accumulation peak is around véraison in seeds and skin (Bogs *et al*., 2005). The last steps of flavan-3-ol monomer biosynthesis are catalyzed by two specific enzymes of the flavonoid pathway: the anthocyanidin reductase (ANR) and the leucoanthocyanidin reductase (LAR). One *ANR* gene and two *LAR* isogenes (*LAR1* and *LAR2*) have been identified in grape (Bogs *et al*., 2005). In young berry, *ANR* and *LAR1* are the most expressed in the two weeks following flowering, synchronously with PA accumulation. In the skin, *ANR* and *LAR2* expression level decreases from two weeks following flowering to véraison, whereas *LAR1* is expressed at very low level. In the seeds, *LAR2* exhibits a peak of expression at véraison, just before the maximal accumulation of PA.

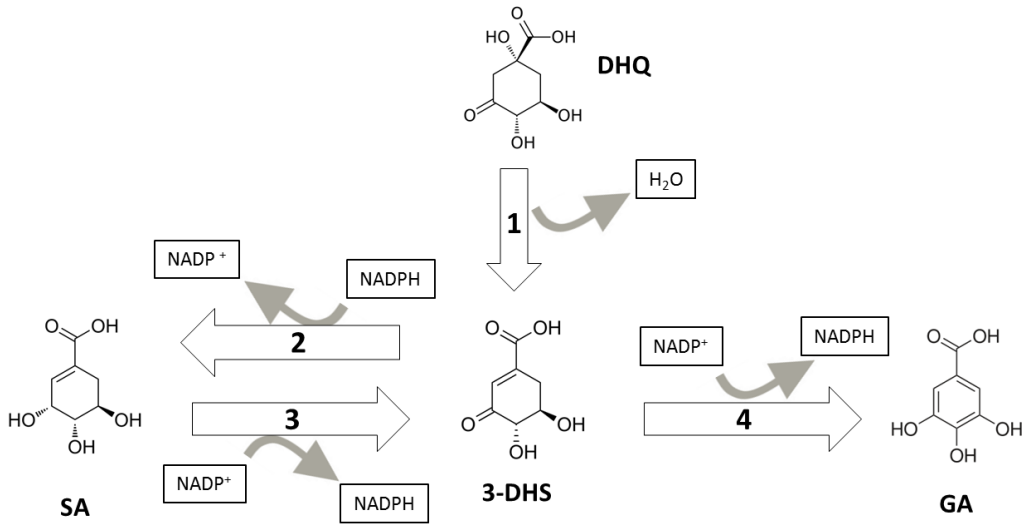


Figure 1.Gallic and shikimic acid biosynthesis in plants.
Reaction 1: Dehydroquinic acid (DHQ) dehydration catalyzed by the dehydroquinate dehydratase (DQD) domain of the dehydroquinate dehydratase/shikimate dehydrogenase (DQD/SDH) producing 3-dehydroshikimic acid (3-DHS). Reaction 2: NADPH-dependent reduction of 3-DHS catalyzed by the shikimate dehydrogenase (SDH) domain of DQD/SDH producing shikimic acid (SA). Reaction 3: $NADP^+$-dependant oxidation of SA catalyzed by the SDH domain of DQD/SDH producing 3-DHS. Reaction 4: $NADP^+$-dependant oxidation of 3-DHS catalyzed by the SDH domain of DQD/SDH producing gallic acid (GA).

Transcriptional factors (TFs) orchestrate the biosynthesis of flavonoids, controlling the expression of the genes involved in the pathway (Winkel-Shirley, 2001). Several TFs controlling positively or negatively PA biosynthesis in grape have been highlighted. Three positive TFs, MybPA1 (Bogs *et al*., 2007), MybPA2 (Terrier *et al*., 2009), MybPAR (Koyama *et al*., 2014), and two negative TFs, MybC2-L1 (Huang *et al*., 2014) and MybC2-L3 (Cavallini *et al*., 2015) have been identified.

The pathway leading to PAs and the network of TFs involved in the regulation of their biosynthesis are well described. However gaps remain to be filled in, notably PA galloylation, transport and storage, and polymerization. Transcriptomic analyses of grapevine hairy-roots overexpressing either MybPA1 or MybPA2 have allowed to establish a list of induced genes and potentially involved in PAs biosynthesis (Terrier *et al*., 2009). A Quantitative Trait Loci (QTL) mapping has been performed to localize regions of the genome able to influence different PAs characteristics, and notably %G (Huang *et al*., 2012). Focusing on PA concentration and galloylation, a restricted list of 20 candidate genes has been determined by combining transcriptomic analyses and QTL mapping (Carrier *et al*., 2013).One gene encoding a dehydroquinate dehydratase/shikimate dehydrogenase (DQD/SDH, here referred as *VvSDH9*) is part of this list.
VvSDH9 is one of the 4 genes encoding DQD/SDH in grapevine genome. A gene (designated *VvSDH5*) is located on chromosome 5. The other 3 genes (designated *VvSDH7, -8* and *-9*) are located side by side as a cluster on chromosome 14, under a QTL of %G.

In plants, DQD/SDH (EC 4.2.1.10/1.1.1.25) is known as a bifunctional enzyme catalyzing the third and the fourth steps of the shikimate (SA) pathway: dehydration of 3-dehydroquinate and NADPH (nicotinamide adenine dinucleotide phosphate)-dependant reduction of 3-dehydroshikimate (3-DHS) (reactions 1 and 2, respectively, Fig. 1). The latter reaction is reversible; SA can be subject to a $NADP^+$-dependant oxidation to form 3-DHS (reaction 3). The SA pathway provides carbon structures used for the biosynthesis of aromatic amino acids and downstream different secondary metabolites, including flavonoids. Otherwise, it has been proposed that some metabolites can be produced *via* branchpoints in the SA pathway, notably quinic acid (Herrmann & Weaver, 1999) and GA (Werner *et al*., 2004). GA is largely found in vascular plants, fruits and foods (e.g. Yilmaz & Toledo, 2004, Lee *et al*., 2008) and is a precursor for the biosynthesis of both galloylated PAs and hydrolysable tannins (Niemetz & Gross, 2005). Its glucose ester, called β-glucogallin (β-G, 1-*O*-β-D-galloylglucose), could be the precursor for galloylation of flavan-3-ols (Liu *et al*., 2012). In plants, GA biosynthesis has been studied since the 1960s (Dewick & Haslam, 1969). A retrobiosynthetic NMR study with labelled glucose supplied to sumac young leaves highlighted that the carboxylic group of GA comes from an intermediate of the SA pathway (Werner *et al*., 1997). From crude extracts of mountain birch leaves (*Betula pubescens* ssp. *Czerepanovii*), GA biosynthesis from 3-DHS and $NADP^+$ has been observed (reaction 4, Fig. 1, Ossipov *et al*., 2003). Recently, it has been demonstrated that, in addition to its already known catalytic properties, DQD/SDH would be responsible for GA biosynthesis in *Juglans regia* (Muir *et al*., 2011).

In the present study, we have cloned and sequenced the 4 genes annotated as DQD/SDH (called *VvSDH*) carried by grapevine genome to test their capacity to produce GA. Heterologous expression of the proteins encoded by *VvSDH* has been monitored in *Escherichia coli*. Kinetic parameters have been determined *in vitro* for SA oxidation and 3-DHS reduction (substrate and cofactor affinity, pH effect) for each VvSDH. *In vitro* GA production from both substrates has been evaluated. *In vivo*, GA, β-G, and %G of PAs have been quantified in grapevine hairy-roots overexpressing *VvSDH8*. Our results suggest that divergences at key amino acids in the SDH domain of DQD/SDH could promote GA biosynthesis from 3-DHS and $NADP^+$.

Material and methods

Plant material

Grape berries (*Vitis vinifera* L. cv. Syrah) were harvested in the INRA-Supagro vineyard (Montpellier, France) at different development stages and classified according elapsed time from anthesis: 11, 18, 35, 52, 56, 58, 65, 86, and 99 days after flowering (daf). Fifty-two daf marks the start of the fruit ripening (véraison). Skin, pulp and seeds of berries harvested at 18 daf (green stage), 52 daf (véraison) and 99 daf (maturity) were isolated. Immediately after harvest, berries were frozen in liquid nitrogen. The fruits were ground using a Dangoumau blender (Dangoumill 300, Lonjumeau, France) and stored at -80°C.

RNA extraction and real-time PCR

Extraction of total RNA from 50 mg of powder was made using RNeasy Plant Mini kit (Qiagen, Hilden, Germany) and following manufacturer instructions. The concentration of total RNA isolated from berries was quantified with Nanodrop 3300 (ThermoScientific, Wilmington, USA). 1 µg of isolated RNA was used for reverse transcription with the kit ImProm-II TM Reverse Transcription System in a total volume of 10 µL (Promega). 5µL of synthesized cDNA were amplified in triplicate by PCR (Polymerase Chain Reaction) using specific primers of *SDH* genes (Supplemental Table 1) and the Power SYBER-Green PCR Master-kit (Applied Biosystems, Applera France, Courtaboeuf, France). PCR were carried out with the 7300 Real-Time PCR System (Applied Biosystems) and analyzed with 7300 System SDS Software v 1.3.1.
To normalize *VvSDH* expression, *EF1α* was used as reference gene (Terrier *et al*., 2005; Monteiro *et al*., 2013). Relative *SDH* expression was calculated from cycle threshold (Ct) according the formula 2 exp-(Ct_{SDH}- $Ct_{EF1\alpha}$).

Cloning and vectors

VvSDH cDNA were amplified from *Vitis vinifera* L. cv. Syrah with high fidelity polymerase (Advantage–HF 2 PCR kit, Clontech, California, USA).The forward and reverse primers for cloning into pGEMT-Easy (Promega, Madison, Wisconsin, USA) were designed to incorporate the restriction site for BamHI before start codon (except for *VvSDH9*) and NotI after stop codon, respectively (Supplemental Table 1). A partial cDNA of *VvSDH9* was cloned into pGEX-4T2 to produce the protein used in this study, lacking the apparent transit peptide constituted of the first 17 amino acids in the N-terminal region as for NtSDH1 (Ding *et al*., 2007), to avoid formation of inclusion bodies. The purified insert was cloned into pGEMT-Easy and digested with the same restriction enzymes. The DNA fragments obtained were cloned into pGEX-4T2 previously digested with the same restriction enzymes (GE Healthcare, Chalfont St Giles, UK).
VvSDH8 and *VvSDH9* full-length cDNA were in parallel cloned in the pENTR/D-TOPO vector (Invitrogen, Courtaboeuf, France) and transferred in the pH2GW7 vector (Karimi *et al*., 2002) by LR reaction (LR clonase II; Invitrogen). The complete sequence of *VvSDH9* cloned in pH2GW7 was used for sequence analysis.
Competent cells of *Escherichia coli* (*E.coli*) DH5α were used to incorporate and propagate all vectors, except pGEX-4T2 for which BL21 (DE3) cells (Stratagene, La Jolla, California, USA) were used.
The incorporation of *VvSDH* cDNA in vectors was checked by PCR on individual colonies using the amplification primers and by vector sequencing (Sanger sequencing, GATC Biotech, Constance, Germany).

Sequence analysis

Amino acids sequence of plant DQD/SDHs were collected from NCBI public database and Phytozome 10.0. Average amino acids identity was calculated by ClustalW2 multiple sequence alignment (McWilliam *et al*., 2013). Phylogenetic analyses were conducted using MEGA version 6 (Tamura *et al*., 2013) to construct a neighbor-joining tree with default parameters. The presence of a putative chloroplastic transit peptide has been predicted using ChloroP 1.1 (Emanuelsson *et al*., 2006).

Heterologous expression in E.coli

LB (Luria-Bertani) medium with ampicillin (100 µg.mL^{-1}) were inoculated with *E.coli* BL21 cells containing *VvSDH* cDNA fused with Glutathione-*S*-transferase (GST) sequence in pGEX-4T2. The bacteria culture was incubated overnight at 37°C under shaking. After overnight growth, 3mL of the pre-culture was used to inoculate an erlenmeyer containing 120mL of YTA 2X medium with ampicillin (100 µg.mL^{-1}). The Erlenmeyer was incubated at 37°C under shaking and bacteria growth was checked following OD_{600} with a spectrophotometer. Once $OD_{600}$was between 0.8 and 1, 0.1 mM final IPTG (isopropyl-1-thio-β-D-galactopyranoside) was added to trigger the induction of GST-SDH fusion proteins. The Erlenmeyer was placed at room temperature under slow shaking during about 20 hours. After this induction step, culture medium was centrifuged (8000g, 10 min, 4°C). The pellet was resuspended in cold PBS (phosphate buffer saline), 0.1% (v/v) β-mercaptoethanol, 1% (v/v) Triton X-100, and lyzosyme (final concentration 1 mg.mL^{-1}). The solution was sonicated in cold conditions to break cell membranes and then centrifuged (12,000 g, 5 min,10°C) to remove the pellet. After adding Glutathione-Sepharose^{TM}4B (GE Healthcare, Chalfont St Giles, UK), the supernatant was set to slow shake on rotary axis for 1 hour at room temperature. After centrifugation (1 500g, 5 min, 10°C), the pellet was resuspended in PBS with 0.1% (v/v) *β*-mercaptoethanol to wash the Glutathione-Sepharose 4B beads. Two supplementary washes were performed under the same conditions before adding a thrombin solution to cleave the GST tag and liberate VvSDH proteins. Proteins were assayed according to Bradford method (Bradford, 1976) using a bovine serum albumin reference curve. Purity and molecular weight of the recombinant proteins were checked by running an aliquot on SDS-Polyacrylamide gel (14%) by electrophoresis (data not shown). Recombinant proteins were diluted in glycerol 16 % (v/v), aliquoted, and stored at -20°C.

Measurement of recombinant proteins activity

For kinetic parameter measurements, VvSDH activity from SA as substrate and $NADP^+$ as cofactor (reaction 3, Fig. 1) was determined following NADPH variation at OD_{340} in UV-cuvettes during 2 minutes with a spectrophotometer (Cary 100 Bio UV-Visible, Varian). The same method was used to measure VvSDH activity from 3-DHS as substrate and NADPH as cofactor (reaction 2, Fig. 1).
All reactions were performed at 30°C, in a final volume of 200 mL containing 100mM Bis-Tris Propane HCl buffer, and started from the adding of 100 ng of enzyme in the reaction mixture. No OD_{340} variation was observed without enzyme.
The effect of pH was assayed with 10 mM of substrate and 2 mM of cofactor using 100 ng of recombinant proteins in 100 mM Bis-Tris Propane buffer at pH 6.5, 7, 7.5, 8, 8.5 and 9.
Michaelis constant for SA ($K_{m\,(SA)}$) was determined at pH 7 and pH 9, with reaction mixture containing 2 mM $NADP^+$ and variable concentration of SA. $K_{m\,(NADP+)}$ was determined at the same pH with 4 mM SA and variable concentration of $NADP^+$. $K_{m\ (3\text{-}DHS)}$ was determined at pH 7 for VvSDH5 and VvSDH8 using 4 mM 3-DHS and NADPH as cofactor (0.4 mM).

For each substrate or cofactor concentration, OD_{340} was plotted as a function of time to determine the slope at the origin. Initial velocity (V_i) was calculated with the NADPH molar extinction coefficient (ε=6.18 x 10^3mol.L^{-1}). V_i data was fitted with Michaelis-Menten hyperbola to yield $K_m$and V_{max} values (Hyper32 program, Version 1.0.0, 2003). The turnover number (k_{cat}) was calculated taking into account the quantity of enzymes introduced in the reactional volume. The catalytic efficiency was calculated as k_{cat}/K_m.

Analysis of enzyme products

For 3-DHS and GA assay, 1 µg of SDH was added to a reaction mixture containing 100mM Bis-Tris Propane HCl buffer pH 9, 4 mM SA or 3-DHS, 2 mM $NADP^+$ and 8 mM ascorbic acid in a final volume of 80 µL. After 2 hours of incubation at 30°C, the reaction was stopped with 20 µL HCl 11.7 N.
5 µL of the reactional mixture was injected on an Agilent 1100 LC system (Agilent Technologies, Waldbronn, Germany) equipped with a diode-array detector (DAD). Compounds were separated using an Atlantis dc18 (250 x 2.1, 5 µm) analytical column (Waters, Milford, MA) protected by an Atlantis dc18 (10 x 2.1, 5 µm) precolumn (Waters, Milford, MA).
The method were developed to separate SA, 3-DHS, $NADP^+$, GA and protocatechuic acid (PCA) using UV-DAD (λ=280nm). Mobile phase consisted of water/formic acid (99.9/0.1, v/v) (eluent A) and methanol/formic acid containing 0.5 % formic acid (99.5/0.5, v/v) (eluent B).
The elution program had a 0.25 mL/min flow rate and started with 0% B. The linear gradients were from 0 to 40% B (0-11 min), from 40 to 50% B (11-15 min) and from 50 to 100% B (15-25 min), followed by washing and reconditioning of the column.

Identification of 3-DHS and GA in enzymatic assays was operated on Ultra Performance Liquid Chromatography-Diode Array Detector-Mass spectrometry (UPLC-DAD-MS)system. Separations were performed using a Waters Acquity UPLC-DAD system (Milford, MA), on a (15 × 1 mm i.d.) Acquity BEH C18 column (Waters, Milford, MA; 1.7 µm), operated at 35°C. Mobile phase consisted of water/formic acid (99/1, v/v) (eluent A) and methanol/formic acid (99/1, v/v) (eluent B). Flow rate was 0.08 mL.min^{-1}.
The elution program was as follows: isocratic for 4 min with 1% B, 1-20% B (4-10 min), 20-98% B (10-12 min), isocratic with 98% B (12-17 min).
ESI-MS/MS analyses were performed with a Bruker Daltonics Amazon (Bremen, Germany) mass spectrometer equipped with an electrospray source and an ion trap mass analyser. The spectrometer was operated in the positive and negative ion mode (end plate off set: -500V; temperature, 200°C; nebulizer gas: 10psi and dry gas, 5L.min^{-1}; capillary voltage, 2.5kV in the positive ion mode, 4.5kV in the negative ion mode). Collision energy for fragmentation used for MS2 experiments were set at 1.
Pure commercial solutions of SA, 3-DHS and GA were used as standards.

Grapevine hairy-roots generation

Agrobacterium rhizogenes strain A4 (Collection Française des Bactéries Phytopathogènes; http://www-intranet.angers.inra.fr/cfbp/) transformed with pH2GW7-VvSDH8 construct were spread over a solid MGL/B spectinomycin dish, and then placed 2-3 days at 28°C. One handle of agrobacterium grown on Petri dish was used to inoculate 5 mL of liquid MS/2 containing vanillin (600 µM). The amount of cells in the inoculum was measured at 600 nm with a spectrophotometer and adjusted to 0.3 absorbance units. *In vitro* growngrapevine seedlings (Macabeu cultivar) had a

minimum of 5 nodes to be inoculated. The caulinar apex and the last 2 emerged leaves were detached. The released limbs were pinched during 2-3 seconds with sterile long forceps previously soaked in the agrobacteria solution. The tube containing the inoculated seedlings was locked and sealed with Parafilm to conserve moisture and placed in a culture chamber (24°C,15h photoperiod).
The generated calli were transplanted on Petri dishes containing LGO medium (Torregrosa & Bouquet, 1997). Each root appearing from calli was transplanted on LGO medium Petri dish and treated as an independent transformation event. The hairy-roots were maintained in culture through monthly subcultures.
After several subcultures, hairy-roots were extracted from agar medium, cleaned, weighted and frozen in liquid nitrogen. The samples were cold-milled using a 6770 Freezer/Mill®cryogenic mill (SPEX® SamplePrep, USA) into a fine powder. The powder was used for DNA and RNA extraction and metabolites analysis. Hairy-root devoid of transgene was used as control hairy-root.

Plant metabolites profiling by UHPLC-QqQ-MS analysis

Approximately 100 mg of frozen powder of grapevine hairy-root or berry tissues were weighed and transferred into a Precellys® tube (7 mL) containing beads. 500 µL of pure methanol were added directly to dilute the powder and mix by vortexing. 3.5 mL of acetone:water:trifluoroacetic acid (70:30:0.05, v/v/v) were then added to the extract. The sample grinding was carried out with a grinder-homogenizer (Precellys24®) following a cycle of 3 sequences of 40s at 5,000 rpm spaced out by 40s. This cycle was repeated 3 times. 3.5 mL of homogenate were transferred to a 15 mL tube and centrifuged (5 min, 4,500 rpm, 4°C). Two fractions of 1 mL of supernatant were dried (Genevac, SP Scientific) and then stored at -80°C.
One of the 2 fractions was used for a direct injection, the other one underwent a phloroglucinolysis step before injection.The phloroglucinolysis reagent was prepared with 0.25 g of phloroglucinol and 0.05 g of ascorbic acid with 5 mL of acidified methanol (final HCl concentration: 0.2 M). The pellet used for direct injection was resuspended in 500 µL of methanol:water:formic acid (50:50:1, v/v/v) and placed in an ultrasonic bath. After 30 min, the sample was centrifuged (15 min, 15,000 rpm, 4°C) andthe supernatant was recovered for analysis.
The pellet used for phloroglucinolysis was resuspended in 700 µL of phloroglucinolysis reagent using a sterile needle. After vortexing, the sample was placed in a water bath at 50°C for PA depolymerisation. After 20 min of incubation, the tube was cooled in ice and 700 µL of ammonium formate buffer 200 mM were added to stop the reaction. After centrifugation (15 min, 15,000 rpm, 4°C), the supernatant was recovered and filtered before analysis.
Hairy-roots and grape berry metabolites were analyzed using Ultra High Performance Liquid Chromatography coupled to triple-quadrupole Mass Spectrometry (UHPLC-QqQ-MS) following a method adapted from that described by Lambert *et al*. (2015). This method has been used for the quantification of GA, β-G after direct injection, and depolymerized PAs subunits after depolymerisation by phloroglucinolysis.
The list of quantified molecules and MRM parameters is available in Supplemental Table 2.

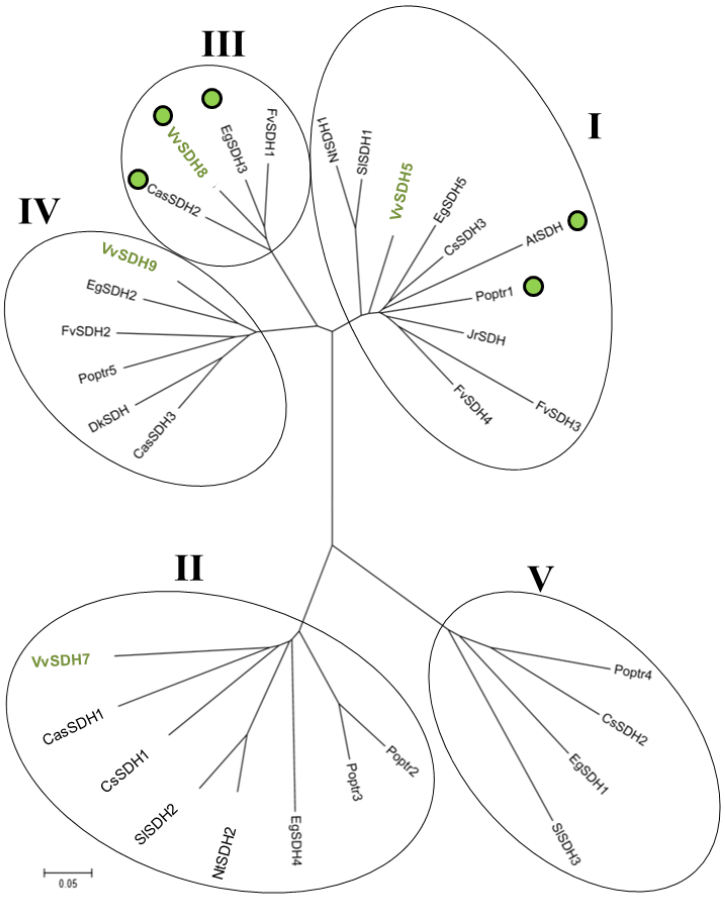


Figure 2. Neighbor-joining tree of selected dehydroquinate dehydratase/shikimate dehydrogenases from Dicots.

This phylogenetic tree was constructed from the 4 VvSDHs sequenced in this study and 28 sequences available on public databases (NCBI and Phytozome) using MEGA6 software. The scale bar represents 0.05 substitutions per site.

Abbreviations are: At (*Arabidopsisthaliana*), Cs (*Citrus sinensis*), Cas (*Camelliasinensis*), Dk (*Diospyros kaki*), Eg (*Eucalyptus grandis*), Fv (*Fragaria vesca*), Jr (*Juglans regia*), Nt (*Nicotiana tabacum*), Poptr (*Populus trichocarpa*), Sl (*Solanum lycopersicum*), Vv (*Vitis vinifera*).

Accession numbers are: **Group I**: AtSDH (AAF08579), FvSDH3 (XP_004289250), FvSDH4 (XP_004288087), VvSDH5, EgSDH5 (Eucgr.J00263.6), NtSDH1 (AAS90325), Poptr1 (Potri.010G019000); SlSDH1 (AAC17991), CsSDH3 (orange1.1g007151m), JrSHD (AAW65140); **Group II**: CasSDH1 (AIZ93902), CsSDH1 (orange1.1g010050m), NtSDH2 (AAS90324), VvSDH7, EgSDH4 (Eucgr.B01770.2), Poptr2 (Potri.013G029900), Poptr3 (Potri.005G043400), SlSDH2 (XP_010327280); **Group III**: EgSDH3 (Eucgr.H04427.1),FvSDH1 (XP_004302480), VvSDH8, CasSDH2 (AJA40947); **Group IV**: EgSDH2 (Eucgr.H04428.1),VvSDH9, Poptr5 (Potri.013G029800), FvSDH2 (XP_004302479), DkSDH (BAI40147), CasSDH3 (AJA40948); **Group V**: Poptr4 (Potri.014G135500), EgSDH1 (Eucgr.H01214.1), CsSDH2 (orange1.1g010101m), SlSDH3 (XP_004242317). Accession number corresponds to NCBI or Phytozome 10.1.2 database. VvSDHs are in green bold letters. The presence of a putative chloroplastic target peptide according to ChloroP is reported by a circle.

Statistical analysis

The significance of the results from HPLC and UHPLC-QqQ-MS analysis has been statistically assessed with a Student's t-test using two-sided alternative with the R software (version 3.12, R Development Core Team, 2009).

Results

Identification of four DQD/SDHs

According to the 12X version of *Vitis vinifera* genome and V2 annotation (http://genomes.cribi.unipd.it/grape/), *VvSDH5* is located on chromosome 5 whereas the 3 other genes are clustered on chromosome 14 (Supplemental Figure 1). This cluster is 34.149 kb long. The genomic DNAs of VIT_205s0020g02030 (*VvSDH5*), VIT_214s0030g00670 (*VvSDH7*), VIT_214s0030g00660 (*VvSDH8*) and VIT_214s0030g00650 (*VvSH9*) are 32.141, 9.695, 4.23 and 9.573 kb long, respectively. The exon-intron pattern corresponding to the cloned cDNAs is shown in Supplemental Figure 1. Full-length cDNAs were amplified from the pericarp of Syrah berries harvested during green stage, cloned and sequenced. *VvSDH5*, *VvSDH7*,*VvSDH8* and *VvSDH9* full-length cDNA were 1560, 1560, 1596 and 1596 nucleotide long, coding for proteins constituted of 520, 520, 532 and 532 amino acids with a predicted molecular weight of 56.54, 57.17, 57.68, and 57.28 kDa respectively (ExPASy server, Gasteiger *et al*., 2003). Full-length cDNAs were used for sequence alignment and phylogenetic analysis. VvSDH5, VvSDH8 and VvSDH9 share more than 70% of nucleotide and amino acid identity. VvSDH7 shares less than 60% of nucleotide identity and about 50% of amino acids identity with the others (Supplemental Figure 2).

Phylogenetic analysis

We have tested if DQD/SDHs from different family species across Dicots could share high sequence identity. The scientific and common name, the family and the number of protein sequences analyzed are listed in Supplemental Table 3.

The neighbor-joining tree has been constructed with *VvSDH* sequences and characterized plant DQD/SDHs from *Arabidopsis thaliana* (*At*, Singh & Christendat, 2006), *Populus trichocarpa* (*Poptr*, Guo *et al*., 2014), *Juglans regia* (*Jr*, Muir *et al*., 2011) and *Nicotiana tabacum* (*Nt*, Ding *et al*., 2007). Considering that GA takes part in tannin metabolism, we have included DQD/SDH sequences of plants known to accumulate high amount of galloylated PAs: *Camellia sinensis* (Cas, Jiang *et al*., 2013), *Diospyros kaki* (*Dk*, Ikegami *et al*., 2005), and/or ellagitannins, that are polyphenols formed from the oxidative linkage of galloyl groups from pentagalloyl glucose, such as *Fragaria vesca* (*Fv*, Landete, 2011) and *Eucalyptus* sp. (*Eg*, Boulekbache-Makhlouf *et al*., 2010). We also choose 2 other plants whose genome has been entirely sequenced, *Solanum lycopersicum* (*Sl*), *Citrus sinensis* (*Cs*), to collect an exhaustive list of DQD/SDH sequences from other Dicots.

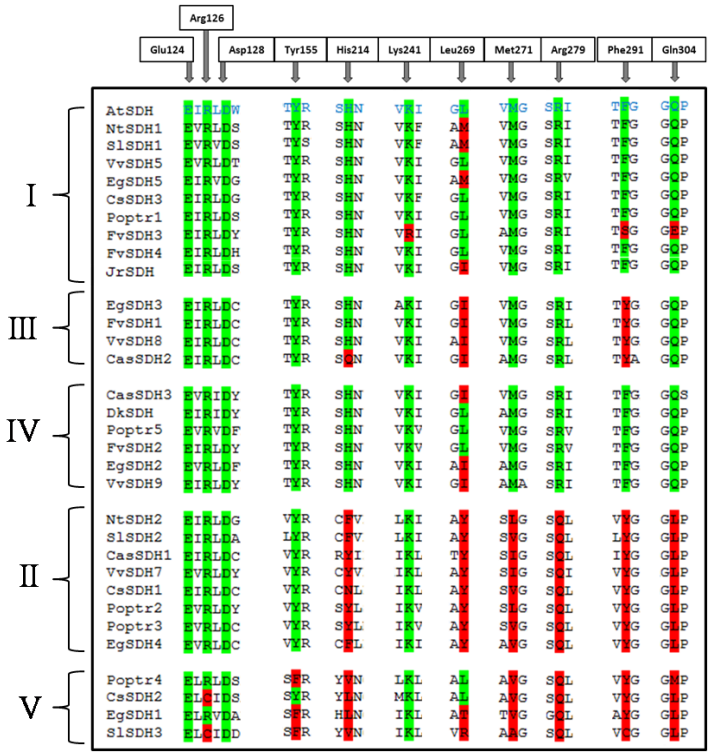
We observed that DQD/SDHs, independently of their species of origin, cluster into5 groups (I-V, Fig. 2).

The 10 proteins constituting group I share about 75% of sequence identity on average. VvSDH5 clusters with 4 characterized DQD/SDHs: AtSDH (Singh & Christendat, 2006), Poptr1 (Guo *et al*., 2014), JrSDH (Muir *et al*., 2011) and NtSDH1 (Ding *et al*., 2007).

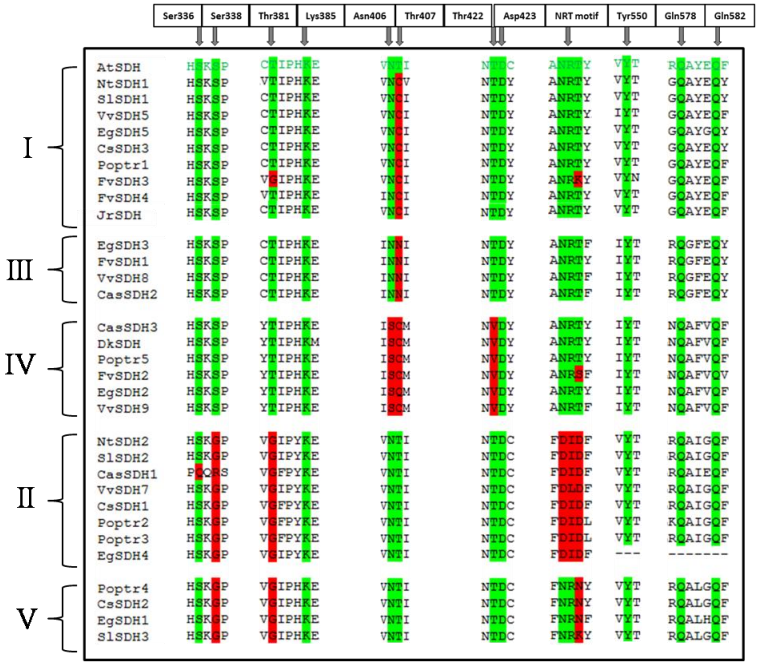
In group II, 8 DQD/SDHs proteins share 71% of amino acid identity on average. In the proposed phylogenetic tree, VvSDH7 clusters with Poptr2 and Poptr3, two characterized DQD/SDHs (Guo *et al*., 2014) and NtSDH2 (Ding *et al*., 2007).

Figure 3.Comparison of key amino acids of DQD/SDHs from Dicots.
(**A**) DQD domain. (**B**) SDH domain. Protein alignement has been performed by ClustalW. Amino acids numerotation is based on DQD/SDH from *Arabidopsis thaliana* (AtSDH, Singh & Christendat, 2006). Key amino acids are shaded in green when identical to AtSDH, and in red when different from AtSDH. The groups I-V have been determined by the neighbor-joining tree. This sequence alignement includes DQD/SDHs from *A. thaliana* (AtSDH), *N. tabacum* (NtSDH1 and 2), *Solanum lycopersicum* (SlSDH1–3), *V. vinifera* (VvSDH5, -7, -8 and -9), *E. grandis* (EgSDH1–5), *Citrus sinensis* (CsSDH1–3), *P. trichocarpa* (Poptr1–5), *F. vesca* (FvSDH1–4), *Juglans regia* (JrSDH), *Camellia sinensis* (CasSDH1–3), *D. kaki* (DkSDH).

A



B



In group III, 4 DQD/SDHs exhibit about 85% of sequence identity on average, representing the higher value among the four groups identified. Interestingly, this value is reached althoughsequences belong to 4 different species (VvSDH8, CasSDH2, FvSDH1 and EgSDH3).
Group IV is constituted of 6 DQD/SDHs belonging to 6 different species and sharing 77% of sequence identity on average. VvSDH9 clusters with Poptr5, a DQD/SDH already characterized in poplar.
Four proteins sharing 68% of amino acid identity on average constitute e group V. None of these proteins have been characterized yet.
A putative chloroplastic transit peptide has been detected for some members of groups I (AtSDH, Poptr1, EgSDH1) and III (CasSDH2, VvSDH8, EgSDH3) using ChloroP 1.1 (Fig. 2).

Comparison of key amino acids for DQD and SDH activity

Key amino acids for both DQD and SDH activities have been highlighted from the crystal structure of AtDQD/SDH (Singh & Christendat, 2006). DQD/SDHs sequences described above have been aligned with AtDQD/SDH to examine the identity at key amino acids in the DQD domain (Fig. 3, A) and the SDH domain (Fig. 3, B).
The first 316 amino acids of AtDQD/SDH form its N-terminal DQD domain. This part of the protein ensures the dehydration of 3-dehydroquinate to produce 3-DHS. Lys 241 (according to AtDQD/SDH numerotation) forms a Schiff base intermediate and His 214 acts as a catalytic base. Lys 241 is well conserved among the examined sequences (except FvSDH3). His 214 is replaced by a Tyr or a Phe in the DQD/SDHs clustering in group II (except CsSDH1, Asn instead of His), and by a Val or a Leu in group V.
Additional residues are involved in the activity of the DQD domain. Leu 269 and Met 271 are localized in the vicinity of the substrate (Singh & Christendat, 2006). We observed a high divergence for Leu 269 that is replaced by a Tyr in group II, by a Met for several DQD/SDHs clustering in group I, and by an Ile in group III and several members of group IV. In group V, EgSDH1 exhibits a Thr and SlSDH3 has an Arg instead of Leu 269. We also observed divergences for Met 271 in groups II and V. Met 271 is substituted by an Ile (CasSDH1 and VvSDH7), a Leu (NtSDH2 and Poptr2) or a Val in the other proteins of group II. In group V, Met 271 is replaced by a Val (except by an Ala in SlSDH3). However, Met 271 is fully conserved in the other 3 groups. Arg 279 has been identified as a key binding group in AtDQD/SDH. DQD/SDHs clustering in groups II and V exhibit a Gln instead of Arg 279. Phe 291 stacks against the carbon backbone of the substrate. A Tyr has been found instead of Phe 291 in the sequence of all members of groups II and III, and in most members of group V. Gln 304, found in the vicinity of the substrate, is replaced by a Leu in the sequence of the members of groups II and V (except Poptr4). Otherwise, four other amino acids shown to be involved in the hydrogen bonding network (Glu 124, Arg 126, Asp 128 and Tyr 155) are highly conserved among the aligned sequences. A Cys instead of Arg126 and a Phe instead of Tyr 155 has been detected in some sequences of group V only.
We also observed amino acid divergences among the SDH domain including the sequence between Ile 328 and Gly 588 of AtDQD/SDH (Fig. 3, B). SDH domain ensures the reversible dehydrogenation of 3-DHS to produce SA. According to Singh & Christendat (2006), Ser 336, Ser 338, and Tyr 550 bind the C1 carboxylate in a trigonal arrangement. Even if Ser 336 and Tyr 550 are well conserved among all sequences, Ser 338 is replaced by a Gly in DQD/SDHs clustering in groups II and V (except an Arg for CasSDH1). The C4 hydroxyl group of the substrate is bound by Asn 406, and the C5 hydroxyl group by Gln 578 and Gln 582. Thr 407 and Thr 422 are involved in substrate orientation. Both Gln residues are fully conserved among the considered DQD/SDHs sequences. However, Asn 406 is

Figure 4. *VvSDHs* expression pattern in grape berry.
The relative expression level of *VvSDHs* was measured along grape berry pericarp development **(A)** and in different tissues (skin, pulp and seed) at three development stages **(B)**.In **B**, grey bars correspond to skin, white to pulp and black to seeds. *Véraison* is marked by the arrow. Gene expression was determined by real-time PCR and normalized with the expression of the reference gene *Elongation factor 1 α* (*EF1α*). Data represent the mean of three replicates. Error bars indicate ± standard deviation.

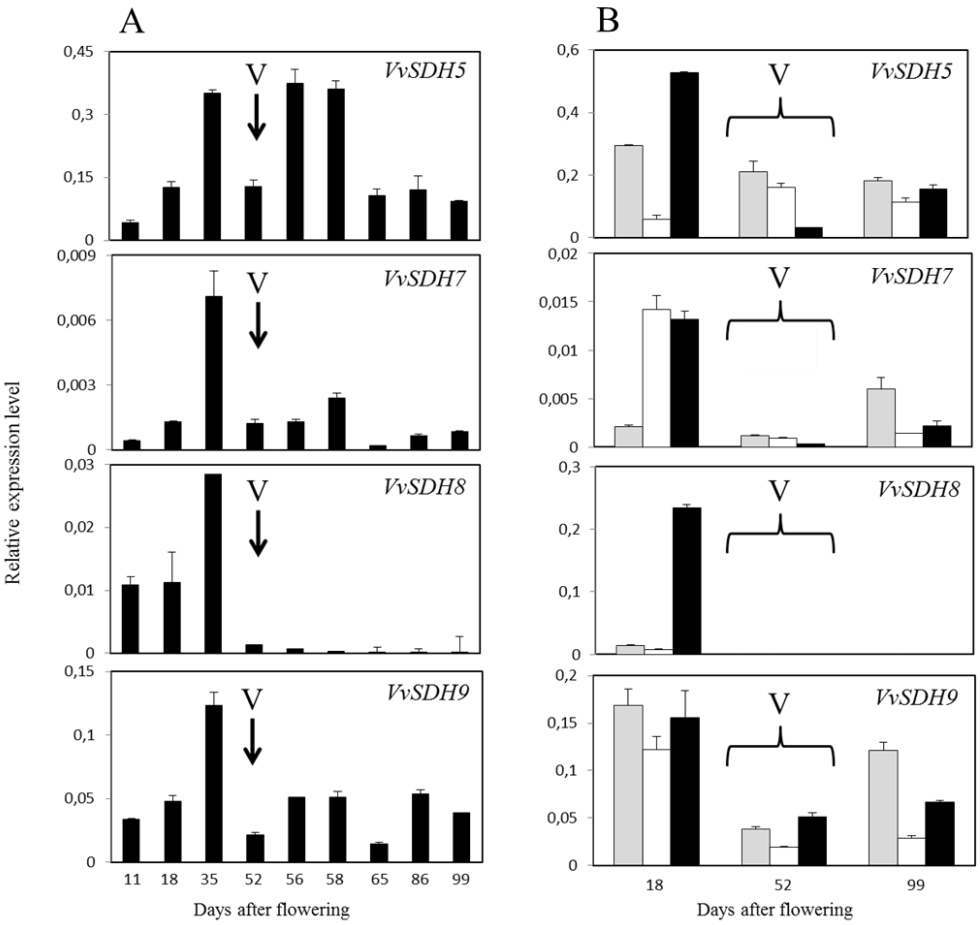
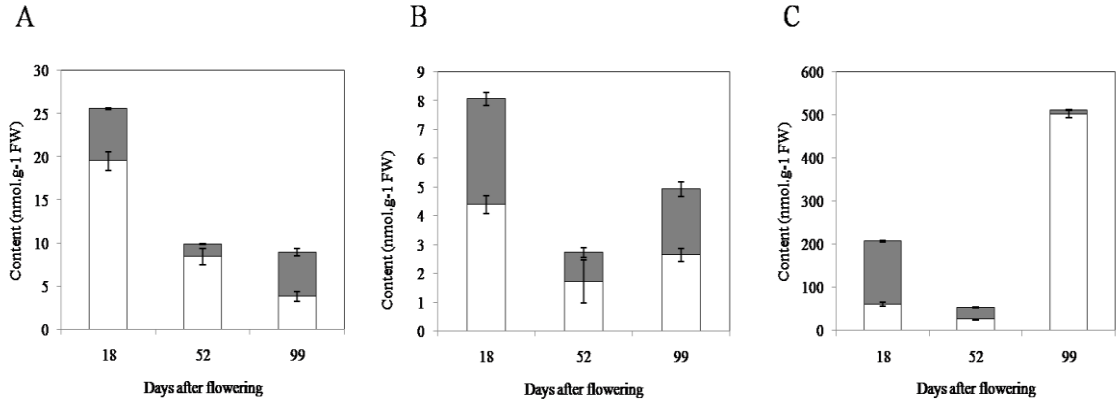


Figure 5. Gallic acid and β-glucogallin contents in grape berry tissues along development.
(A) Skin. (B) Pulp. (C). Seeds.White and grey bars correspond to GA and β-glucogallin, respectively. Data represent the mean of three replicates. Error bars indicate ± standard deviation. Samples corresponding green stage, *véraison* and maturity were collected at 18, 52 and 99 days after flowering, respectively. FW: fresh weight.



replaced by a Ser in the members of group IV. Thr 407 is replaced by a Cys in groups I (except AtSDH) and IV (except EgSDH2), and by an Asn in group III. Moreover, Thr 422 is replaced by a Val in group IV. Lys 385 and Asp 423 were expected to be catalytic residues due to their proximity with the C3 hydroxyl, the site of deprotonation. Those two amino acids are fully conserved among the analyzed sequences. Thr 381 has been highlighted as an important residue for substrate selectivity, i.e. shikimic or quinic acid (Singh & Christendat, 2006). We found a Gly instead of Thr 381 in the sequences of groups II and V, and in FvSDH3 sequence (group I).
DQD/SDHs carry a cofactor-binding motif of 3 successive amino acids that correspond to Asn 484, Arg 485 and Thr 486 in AtDQD/SDH sequence. The presence of Arg 485 in this NRT motif attests NADP(H) binding (Singh *et al*., 2008). The presence of Asp instead of Arg in the NRT motif is believed to be characteristic of NAD(H)-binding shikimate dehydrogenases (Michel *et al*., 2003). We noted that only DQD/SDHs from group II exhibit an Asp in the NRT motif.

VvSDHsexpression pattern, GA and β-G content in grape berry

The expression pattern of each *VvSDH* in the pericarp has been determined along grape development *via* quantitative real time PCR (Fig. 4, A). *VvSDH7* and *VvSDH9* expression is relatively stable during grape development, except a peak of expression before véraison (35 daf). High expression is also observed for *VvSDH5* before véraison (35 daf) but its maximum expression level is reached after véraison (56 daf). Interestingly, *VvSDH8* expression level rises during the green stage to reach a peak at 35 daf. At véraison, its expression drops and stays very low throughout the ripening stage.
VvSDH expression has been also analyzed in the main berry tissues (skin, pulp and seeds) at three development stages (green stage, véraison and maturity) to determine their spatial and temporal expression pattern (Fig. 4, B). Generally, the expression level of*VvSDH5*, *-7* and *-9* is maximal during the green stage (18 daf), low at véraison and intermediate atmaturity. During the green stage, *VvSDH5* is less expressed in the pulp than in skin and seeds, *VvSDH7* is less expressed in skin than in pulp and seeds, whereas *SDH9* expression is relatively high in the three tissues. *VvSDH8* has a very low level of expression whatever the tissue or the developmental stage, except in the immature seeds.
The content of GA and its glucose ester, β-G, was measured in skin, pulp and seeds, at 3 development stages: green stage (18 daf), véraison (52 daf) and maturity (99 daf) (Fig. 5). At the same development stage, GA and β-G contents were always higher in seeds than in pulp and skin. The molar ratio β-G/GA was systematically lower or close to 1, except in immature seeds in which this ratio was close to 2.4. In skin and pulp, GA content was maximal during the green stage (~20 and 4 µmol.g^{-1} FW, respectively, Fig. 5,A& B). As a result, GA must be synthesized before véraison. The highest content measured in mature seeds (~500 nmol.g^{-1} FW, Fig. 5 C) could be due to the release of galloyl moieties from galloylated flavan-3-ols, as their content decreases from véraison in seeds (Kennedy *et al*., 2000, Downey *et al*., 2003).

Enzymatic activity of the recombinant enzymes

GST-tagged VvSDHs have been produced in *E.coli*. After purification and tag removal, the recombinant proteins have been tested *in vitro*. NADPH-dependent reduction of 3-DHS (reaction 2, Fig. 1) and NADP$^+$-dependent oxidation of SA (reaction 3) have been tested.
The influence of pH buffer on enzymatic assays has been tested. Kinetic parameters (K_m, V_{max}) were determined at pH 7 to mimic cytosolic conditions and at pH 9, a more alkaline condition closer to the pH of chloroplast stroma (Song *et al*., 2004), using 100 ng of proteins. Actually, even if the SA pathway is considered as chloroplastic, NtSDH2 has been localized in the cytosolic compartment in *Nicotiana tabacum* (Ding *et al*., 2007). Very low VvSDH activity under certain conditions did not

Figure 6. pH effect on VvSDH specific activity.
VvSDH activity was measured following NADPH variation by spectrophotometry during 2 minutes after adding enzyme. The Bis-Tris Propane HCl buffer allowed measurements at pH6.5, 7, 7.5, 8, 8.5 and 9. The reaction mixture contained 10 mM SA, 2 mM $NADP^+$ to test reaction 3 described in Fig. 1 (A) or 2 mM 3-DHS and 0.4 mM NADPH to test reaction 2 described in Fig. 1 (B). VvSDH7 data for reaction 3 are presented independently due to its lower activity. Data represent the mean of three replicates. Error bars indicate ± standard deviation.

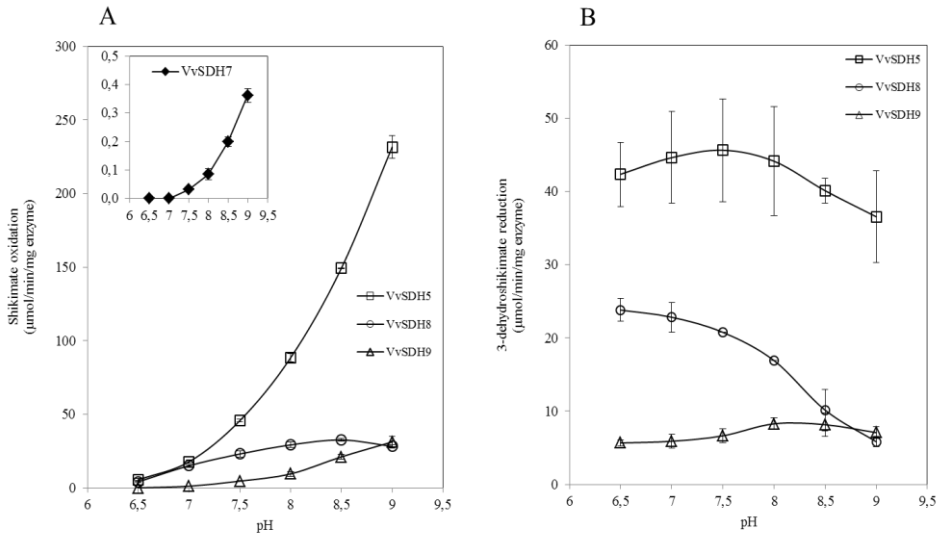


Table 1. Michaelis-Menten kinetic parameters of recombinant VvSDHs.

Enzyme	Molecule	pH	K_m (µM)	V_{max} (µkat.mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}.s^{-1}$)
VvSDH5	$NADP^+$	9	214.7 ± 56.4	4.23±0.47	239 ± 26.3	1113
	$NADP^+$	7	117.2 ± 28.3	0.54±0.05	30.70 ± 3.1	262
	SA	9	90 ± 16	1.67±0.13	94.55 ± 7.2	1051
	SA	7	155 ± 32.6	0.21±0.01	11.72 ± 0.77	76
	3-DHS	7	156.9 ± 31.3	1.70±0.18	95.85 ± 10.19	611
VvSDH7	$NADP^+$	9	nd			
	$NADP^+$	7	nd			
	SA	9	nd			
	SA	7	nd			
	3-DHS	7	nd			
VvSDH8	$NADP^+$	9	38.4 ± 11.9	0.45 ±0.05	26.11 ±2.68	680
	$NADP^+$	7	23 ± 4.9	0.34 ±0.02	19.78 ±1.28	862
	SA	9	27.3 ± 10.4	0.09 ±0.01	5.16 ±0.56	189
	SA	7	422.4 ± 88.3	0.13 ±0.01	7.56 ±0.65	18
	3-DHS	7	46.3 ± 18.2	0.15 ±0.02	8.62 ±0.93	186
VvSDH9	$NADP^+$	9	53.8 ± 5.8	0.50 ±0.02	28.89 ±1.06	537
	$NADP^+$	7	nd			
	SA	9	146.8 ±60	0.46 ±0.09	26.12 ±4.97	178
	SA	7	nd			
	3-DHS	7	nd			

allow us to determine kinetic parameters. The optimal pH for each enzyme has been determined using a pH range buffer from pH 6.5 to 9.
For SA oxidation, VvSDH5 and -9 specific activity increasedwith pH up to 9, whereas VvSDH8 reached its maximum at pH 8.5 (Fig. 6, A). No activity has been detected at pH 7 for VvSDH7 and VvSDH9. For 3-DHS reduction (Fig. 6, B), VvSDH8 exhibited a maximal specific activity at pH 6.5 andthe activity decreased when the pH became more alkaline. VvSDH5 and VvSDH9 reached their maximal specific activity at pH 7.5 and 8, respectively.
Substrate and cofactor kinetic parameters have been determined for reaction 2 (3-DHS reduction) and 3 (SA oxidation) at pH 7 and 9, when it was possible. Kinetic parameters could not be measured for VvSDH7 due to its low activity.
Considering enzymatic SA oxidation at pH 9, VvSDH8 had the higher affinity for SA ($K_{m\ (SA)}$= 27.3 µM), compared to VvSDH5 ($K_{m\ (SA)}$= 90 µM) and VvSDH9 ($K_{m\ (SA)}$= 146.8 µM, Table 1).VvSDH5 had the highest affinity for SA at pH 7. $K_{m\ (SA)}$ values are close to those measured for*Nicotiana tabacum* NtDQD/SDH1 (Ding *et al*., 2007). VvSDH8 and VvSDH9 presented a similar catalytic efficiency (k_{cat}/K_m) at pH 9 for SA, but this value was about 6 times lower than VvSDH5 catalytic efficiency under the same conditions. Both VvSDH5 and -8 had more affinity for $NADP^+$ at pH 7 than at pH 9.
At pH 7, VvSDH8 had a higher affinity than VvSDH5 for 3-DHS ($K_{m\ (3\text{-}DHS)}$= 46.3 µM and 156.9 µM, respectively). However, the catalytic efficiency was about 3 times higher for VvSDH5 (k_{cat}/K_m=611) compared to VvSDH8 (k_{cat}/K_m=186). VvSDH9 kinetic parameters could not be determined for the NADPH-dependent reduction of 3-DHS due to its low activity at pH 7.
No activity was detected using GST produced by *E.coli* from empty vector (data not shown).

GA production by recombinant VvSDH

GA biosynthesis by each recombinant VvSDH was monitored from SA or 3-DHS as substrate and $NADP^+$ as cofactor (reaction 4, Fig. 1). The identity of enzymatic products (3-DHS and GA) has been checked by extracted ion chromatogram and fragmentation spectrum compared with injection of corresponding commercial standards (Supplemental Fig. 3&4).
VvSDH specific activity from SA being higher at pH 9 (Fig 7, A), we chose this pH condition to characterize their capacity to produce GA.
From 4 mM SA, 3-DHS and GA were not detected without enzyme in HPLC chromatogram (Fig. 7, A). VvSDH7 produced 3-DHS (7.39 µM) but no GA from SA (Fig. 8, A). VvSDH5, -8 and -9 activity resulted in the formation of both 3-DHS (53.26, 43.99 and 59.89 µM, respectively) and GA (6.26, 5.09 and 1.12 µM, respectively) (Fig 8, A). Concerning GA production, VvSDH5 was the most efficient enzyme from SA.
From 4 mM 3-DHS, a spontaneous formationof protocatechuic acid (PCA) and GA was detected without enzyme (Fig. 7, B), that explains why GA production was also observed from SA in presence of enzyme (Fig. 7, A). HPLC analyzes highlighted the instability of 3-DHS in the enzymatic assay medium (Bis-Tris Propane HCl pH 9). Chemically, PCA formationfrom 3-DHS involves dehydration and enolization, and GA formation occurs through oxidation and enolization (Dewick, 2002). Non-enzymatic formation of PCA from 3-DHS has been reported earlier (Richman *et al*., 1996, Li *et al*., 1999) and has been observed in alkaline conditions (pH 10, Ossipov *et al*., 2003). Their formation increased with pH (data not shown). VvSDH9 and VvSDH8, to a lesser extent, were able to significantly increase GA production compared to the control condition without enzyme (Fig.8, B).Muir and coworkers reported that a DQD/SDH from *Juglans regia* produced about 40 µM 3-DHS and 14 µM of GA from SA, and about 25 µM GA from 3-DHS. The amount of GA obtained in the

Figure 7. Chromatograms of HPLC analysis using UV-DAD (λ = 280 nm) of reaction assay.
(A) Control condition (blue line) contained Bis-Tris Propane pH9 100 mM, 4 mM 3-DHS, 2 mM $NADP^+$, 8 mM ascorbic acid. Enzymatic assay (red line) contains the same mixture and 1 µg VvSDH5 were added to start the reaction. (B**)** Control condition (blue line) contained Bis-Tris Propane pH9 100 mM, 4 mM SA, 2 mM $NADP^+$, 8 mM ascorbic acid. Enzymatic assay (red line) contains the same mixture and 1 µg VvSDH9 were added to start the reaction.

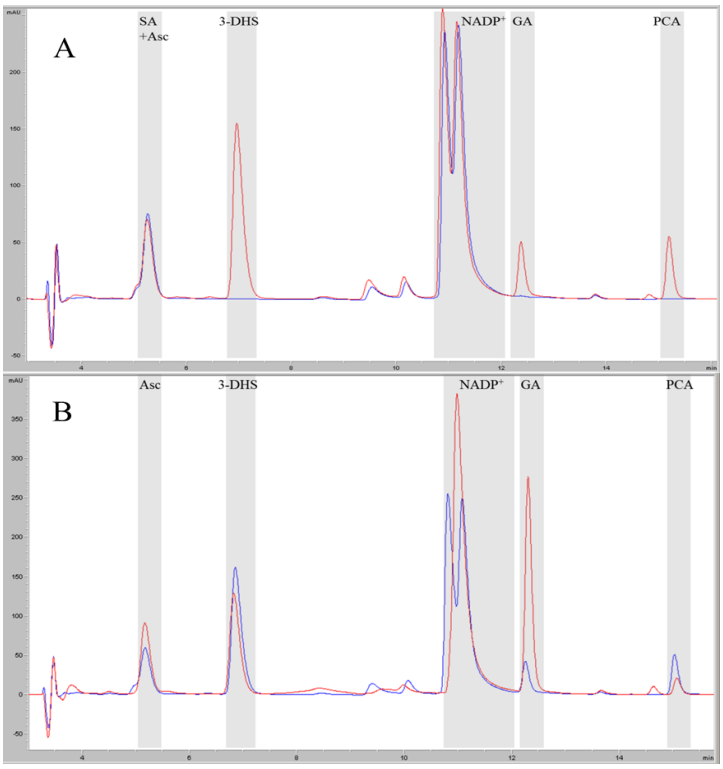
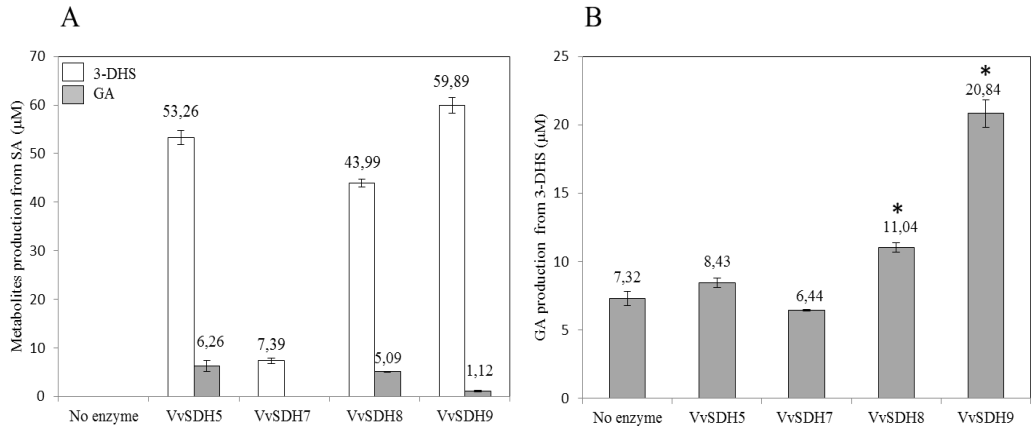


Figure 8. HPLC assay of enzymatic reactions using UV-DAD (λ = 280 nm).

(A) HPLC quantification of 3-DHS (white bar) and GA (grey bar) in the condition lacking enzyme and in the presence of each VvSDH *in vitro* from SA, at pH9. (B) GA quantification in the condition lacking enzyme and in the presence of each VvSDH *in vitro* from 3-DHS, at pH9. Each bar represents the mean value of three assays ± standard deviation. The significance of the results was statistically assessed with a Student's t-test using two-sided alternative. *: $P<0.01$.



same reaction conditions are quite similar but after a much shorter reaction time (2h here instead of 16h).GA production by VvSDH5 and -7 was not significantly different from the control condition lacking enzyme.*In vitro* assays support that GA can be produced from 3-DHS by (i) spontaneous transformation in alkaline conditions and (ii) enzymatic conversion by some of the VvSDHs. Being in substrate saturating conditions, we can consider that those minor spontaneous transformations have not disturbed our experiments.

GA metabolism enhanced by VvSDH8 overexpression in hairy-roots

GA derivatives (free GA, β-G) and flavan-3-ol composition were measured by UHPLC-QqQ-MS analyses in 3 independent lines of hairy-roots transformed with *VvSDH8* (numbered 3A, 6A and 9A) and compared with a control line devoid of transgene. In spite of different tries, we could not obtain transgenic hairy-roots overexpressing *VvSDH9*, and we suspect a detrimental effect of *VvSDH9* overexpression for root development. The relative expression level of *VvSDH8* has been measured in control and transgenic lines, using *VvEF1α* as reference gene (Supplemental Fig. 5).
The lines 3A and 6A tended to accumulate more GA than the control line (+22 and 77%, respectively) and 9A contained significantly more GA than the control line (Table 2). All transformed lines contained significantly more β-G (up to +165 % for line 9A) and galloylated flavan-3-ols (between +80% and +124%) than the control line.
Total content of flavan-3-ols was not significantly different in the lines 3A and 6A transformed with *VvSDH8*. The line 9A contained significantly more flavan-3-ols compared to the control line. %G was significantly higher in the 3 lines overexpressing *VvSDH8* compared to the control line.

Discussion

VvSDHs exhibit different activities towards SA pathway metabolites in vitro

Among the 4 VvSDHs, VvSDH5, -8 and -9 proteins exhibited a "classical" SDH activity from SA, using $NADP^+$ as cofactor. On the contrary, VvSDH7 had very low SDH activity.
Specific activity for SA oxidation reached a maximum at a pH close to 9 for VvSDHs. Similar measurements have been reported with plant DQD/SDHs (Ding *et al*., 2007, Singh & Christendat, 2007). We observed that pH optimum was more acidic for 3-DHS reduction. $K_{m\ (SA)}$ and $K_{m\ (NADP+)}$ values at pH 9 were close to those measured for NtSDH1 (Ding *et al*., 2007). $K_{m\ (3\text{-}DHS)}$ at pH 7 is close to $K_{m\ (SA)}$ at pH 9 for VvSHD5 and VvSDH8, indicating that they have similar affinity for their substrate in both reaction directions. We did not find studies reporting $K_{m\ (3\text{-}DHS)}$ from heterologously produced plant DQD/SDHs, K_m being determined from SA only. However, SDH activity of DQD/SDHs has been observed with purified enzymes from pea (Mousdale *et al*., 1987) and a $K_{m\ (3\text{-}DHS)}$ = 210 µM (with 0.35 mM NADPH, at pH 7) was reported, close to the values obtained in the present work with VvSDHs.
VvSDH5 had a 6 times higher catalytic efficiency from SA at pH 9 compared to VvSDH8 and -9 (Table 1). Concerning GA biosynthesis, VvSDH5 produced the highest amount of GA from SA but GA production was not significantly different from the condition without enzyme with 3-DHS as substrate. Thus, the assayed GA with SA as substrate could result from the spontaneous conversion of the enzymatically formed 3-DHS, rather than enzymatic activity. In our neighbor-joining tree, VvSDH5 clustered notably with the characterized AtDQD/SDH, the unique DQD/SDH encoded in the *Arabidopsis thaliana* genome (Singh & Christendat, 2006), NtSDH1 from tobacco (Ding *et al*., 2007),

Table 2. Gallic acid, β-G and flavan-3-ol content in grapevine hairy-roots.

	Control	**3A**	**6A**	**9A**
Gallic acid	3.78 ± 0.31	4.60 ± 0.33 *	6.70 ± 1.20 *	7.65 ± 0.65 **
β-glucogallin	5.74 ± 0.66	10.57 ± 0.47 **	15.18 ± 1.05 **	8.77 ± 0.38 **
Galloylated flavan-3-ols	0.082 ± 0.002	0.148 ± 0.004 **	0.157 ± 0.005 **	0.184 ± 0.001 **
Total flavan-3-ols	2.87 ± 0.06	3.07 ± 0.06 *	2.97 ± 0.06	3.36 ± 0.08 **
%G flavan-3-ols	2.65 ± 0.01	4.45 ± 0.05 **	4.84 ± 0.07 **	4.93 ± 0.09 **

Gallic acid and β-glucogallin contents are reported as $nmol.g^{-1}$ HR frozen powder. Flavan-3-ols contents are reported as $\mu mol.g^{-1}$ HR.

Galloylated flavan-3-ols content was calculated as the sum of epicatechin gallate and epicatechin gallate with phloroglucinol adduct.

% galloylation (%G) of flavan-3-ols was calculated as a molar ratio.

Each data represents the mean value of 3 assays ± standard deviation. The significance of the results was statistically assessed with a Student's t-test using two-sided alternative. *: $0.01<P<0.05$, **: $P<0.01$.

JrSDH from walnut (Muir *et al*., 2011) and Poptr1 from poplar (Guo *et al*., 2014). Among the key amino acids involved in SDH activity, the only difference detected between AtSDH and VvSDH5 was a Cys instead of Thr 407 (Fig. 3,B). VvSDH5 and Poptr1 had the highest SDH activity in saturating conditions among the DQD/SDHs characterized in their respective genome. During grape development, *VvSDH5* was highly expressed both before and after véraison contrary to the other *VvSDHs*. Even if this gene was mainly expressed in skin and seeds during the green stage, its relative expression was similar in the 3 tissues at véraison and maturity. We can hypothesize that *in planta*, VvSDH5 is devoted to channel the carbon flux towards aromatic amino acids biosynthesis in the main trunk of the SA pathway, converting efficiently 3-DHS to SA whatever the pH and SA to 3-DHS, but more efficiently at higher pH, fitting with chloroplastic conditions. The proteins clustering in group I must be essential for an optimal plant metabolism and development. Indeed, the silencing of NtSDH1 triggered retardation of plant growth and a reduction of aromatic amino acids (Ding *et al*., 2007). This idea is also supported by the fact that all examined species, except *Diospyros kaki* and *Camellia sinensis*, carry a DQD/SDH from group I. However, it is likely that these two species whose complete genome sequencing is not yet available carry other DQD/SDH genes.

VvSDH8 and -9 had lower catalytic efficiency (k_{cat}/V_{max}) for SA than VvSDH5 at pH 9. However, *in vitro* assays showed that these enzymes catalyze GA production. Their SDH activity must be influenced by the redox state in the cellular compartment where they are located, and more precisely $NADP^+$/NADPH ratio. In presence of NADPH, SDH converts efficiently 3-DHS into SA. In presence of $NADP^+$, SDH converts 3-DHS into GA. From the same metabolite (3-DHS), GA production was lower than SA production. This feature is consistent with the fact that enzymes involved in specialized metabolites biosynthesis are less efficient than those involved in primary metabolism (Milo & Last, 2012).

VvSDH8 and -9 clustered in groups III and IV, respectively. VvSDH8 clustered with not yet characterized DQD/SDHs that share the higher average sequence identity of all groups (85.4 %). VvSDH9 clustered with the characterized DQD/SDH Poptr5 from poplar. Both enzymes exhibited "classical" SDH activity from SA and $NADP^+$, but lower than VvSDH5 and Poptr1 (Guo *et al*., 2014). Interestingly, VvSDH9 also clustered with a DQD/SDH identified in persimmon (DkSDH). In non-astringent persimmon fruits, containing low levels of free GA and galloylated PAs, this DkSDH is downregulated compared to astringent persimmon fruits (Akagi *et al*., 2009). The species possessing DQD/SDH clustered in groups III or IV are known to produce high amounts of ellagitannins (strawberry, eucalyptus) or galloylated PAs(tea plant, grape, persimmon). In contrary, the species devoid of DQD/SDHs from group III and IV, such as Arabidopsis, tobacco, tomato and *Citrus sinensis*, do not produce galloylated tannins.

It has been postulated that the C-terminal SDH domain of the bifunctional enzyme DQD/SDH of plants could be involved in GA biosynthesis (Muir *et al*., 2011). Microbial SDH activity is ensured by the monofunctional AroE enzyme and is sufficient to produce GA, as demonstrated for *Escherichia coli*. Proteins of group III (including VvSDH8) present an Asn instead of the Thr 407 identified in AtSDH, or Cys for the other group I members. Compared to the group I proteins, the DQD/SDHs from group IV (including VvSDH9) present a Ser instead of Asn 406 and a Val instead of Thr 422. Substrate orientation must be affected by the substitutions of Thr 407 (VvSDH8) or Thr 422 (VvSDH9). Moreover, substrate binding at C4 position must be influenced by the substitution of Asn 406 (VvSDH9). Those divergences could affect the hydrogen-bonding network andexplain why VvSDH8 and -9 acquired the capacity to produce GA. Modification of substrate orientation might

favor an oxidation of the C5 hydroxyl of 3-DHS, necessary to form 3,5-DHS, that is spontaneously converted into GA (Muir *et al*., 2011).
The relative expression pattern of *VvSDH8* and -*9*, highest before véraison, was synchronous with GA biosynthesis in grape berry.Otherwise, the highest expression of those genes has been detected in the tissues in parallel with the biosynthesis of galloylated PAs (Kennedy *et al*., 2000, Bogs *et al*., 2005): immature seeds for *VvSDH8*, immature skin and seeds for *VvSDH9*. VvSDH8 and -9 probably have a minor role in the SA pathway, due to their low "classical" SDH activity, but probably orient the carbon flux towards GA metabolism. Actually, *VvSDH8* overexpression in grapevine hairy-roots resulted in an increase in GA, β-G and %G in the transformed lines.

VvSDH7 exhibited a very low "classical" SDH activity compared to the other VvSDHs. VvSDH7 clusteredin group II with characterized poplar DQD/SDHs that did not exhibit "classical" SDH activity but rather quinate dehydrogenase activity from quinate and NAD^+ (Poptr2 and -3, Guo *et al*., 2014). Contrary to the others VvSDHs, VvSDH7 carries a Gly instead of Thr 381 in the SDH domain. This residue is determinant for substrate selectivity and this substitution lets think that the protein prefers quinic acid rather than SA (Singh & Christendat, 2006). Similar substitutions have been observed for members of group II (Fig 3,B). Ding and coworkers (2007) have suggested that "the*in vivo* preferred substrate of NtSDH2 is in fact a derivative of shikimate exhibiting a largerfunctional group at the C1 position ring", that corresponds to the description of quinate. Interestingly, this protein has been localized in the cytosolic compartment. The cofactor-binding motif of VvSDH7 exhibits a Asp instead of Arg 484 (Fig 3,B), suggesting that this enzyme is NAD(H)-dependant contrary to the other VvSDHs. Due to the absence of SDH activity and high sequence identity with Poptr2 and -3, VvSDH7 is a good candidate for the NAD^+-dependant biosynthesis of quinic acid. *In vitro* activity of VvSDH7 must be further examined to check if this enzyme exhibits quinate dehydrogenase activity.

How can GA reach the vacuole to take part inflavan-3-ol galloylation?

SA pathway localization in vascular plants has been discussed in the scientific literature. Most of the enzymes constituting this pathway present a plastidic transit peptide that allows their targeting to the chloroplasts (Della-Cioppa *et al*., 1986). However, isoforms devoid of this transit peptide have been identified (e.g. Ganson *et al*., 1986). Moreover, some SA pathway enzymes are active with their transit peptide, indicating that they could have a cytosolic activity (e.g. Schmid *et al*., 1992). Those observations support the existence of a parallel cytosolic SA pathway. Interestingly, two tobacco SDH isoforms have been localized in different compartments (Ding *et al*., 2007). Indeed, NtSDH1 carries a plastid transit peptide and is conventionally found in the chloroplast. NtSDH2 lacks the plastidic targeting sequence and is found in the cytosol. A putative chloroplastic target peptide has been detected for some members of groups I (AtSDH, Poptr1) and III (CasSDH2, VvSDH8, EgSDH3) according to ChloroP analysis (Fig. 3). However, the reliability of this *in silico* analysis is doubtful since it does not predict a chloroplastic target peptide for the chloroplastic NtSDH1. Despite requiring experimental confirmation, members of groups I and III are suspected to be chloroplastic enzymes.
It has been reported that GA is equally distributed in the chloroplast and the vacuole in mesophyll cells protoplasts of *Saxifraga stolonifera* (Taneyama, 1992). However, the mechanisms routing GA towards the vacuole are unknown. Being a weak acid, GA could freely cross the chloroplastic membrane as suggested for *p*-aminobenzoate (Eudes *et al*., 2008). To date, no GA transporter has been identified in plants. The characterization of a chloroplastic transporter of salicylic acid (Serrano *et al*., 2013) suggests that other hydroxybenzoic acids such as GA could be routed to the cytosol *via* a similar mechanism.

About 70% of GA is found as its glucose ester form in immature seeds. Specific grape UDP-glucosyltransferases (VvGTs) expressed in the early steps of seed development are involved in the biosynthesis of β-G, as demonstrated *in vitro* (Khater *et al*., 2012). This type of UDP-Glucosyltransferase activity is known to occur in the cytosolic compartment (e.g. Achnine *et al*., 2005, Pang *et al*., 2013, Mugford *et al*., 2013). If GA is synthesized in chloroplasts, it must be exported to the cytosol to be available for β-G biosynthesis catalysed byVvGTs (Supplemental Figure 6).

This GA glucose-ester is a precursor for flavan-3-ol galloylation in tea plant (Liu *et al*., 2012). The gene responsible for this acylation has never been characterized in plants. In different metabolic pathways, acylation from glucose-esters is catalyzed by Serine Carboxypeptidase Like (SCPL) acyltransferases (reviewed by Bontpart *et al*., 2015). Several studies suggest that SCPL acytransferase(s) could be involved in flavan-3-ol galloylation (Ikegami *et al*., 2007, Terrier *et al*., 2009, Liu *et al*., 2012, Carrier *et al*., 2013). SCPL acytransferase(s) have been localized in the vacuole (e.g. Hause *et al*., 2002, Mugford *et al*., 2009, Nishizaki *et al*., 2013). Hence, flavan-3-ols galloylation could occur in the vacuolar compartment. Two SCP-Like genes (called *VvGAT1* and *-2*) are induced by MybPA1 and MybPA2 according to transcriptomic analyses and localized under a QTL influencing PA %G (Terrier *et al*., 2009, Huang *et al*., 2012). The characterization of those genes could validate the final step of the two-step PAs galloylation that we have hypothesized.

To conclude, we have shown that shikimic and gallic acid metabolisms in grapevine are ensured by different genes belonging to the same gene family. The biochemical characterization and sequence analysis of VvSDH8 and -9 let think that these enzymes could orient 3-DHS *via* its SDH domain towards the GA metabolism in addition to SA pathway. Site-directed mutagenesis would be necessary to check our hypothesis concerning the structure-function relationship of VvSDHs and the involvement of some amino acids in the SDH domain of DQD/SDH producing GA.

This study opens the way for the identification of genes specifically involved in GA production in plants.

References

Achnine L, Huhman DV, Farag MA, Sumner LW, Blount JW, Dixon RA. 2005. Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. The Plant Journal **41**: 875–887.

Akagi T, Ikegami A, Suzuki Y, Yoshida J, Yamada M, Sato A, Yonemori K. 2009. Expression balances of structural genes in shikimate and flavonoid biosynthesis cause a difference in proanthocyanidin accumulation in persimmon (*Diospyros kaki* Thunb.) fruit. Planta **230:** 899-915.

Arnold RA, Noble AC, Singleton VL. 1980. Bitterness and astringency of phenolic fractions in wine. Journal of Agricultural and Food Chemistry **28**: 675-678.

Bogs J, Downey MO, Harvey JS, Ashton AR, Tanner GJ, Robinson SP. 2005. Proanthocyanidin synthesis and expression of genes encoding leucoanthocyanidin reductase and anthocyanidin reductase in developing grape berries and grapevine leaves.Plant Physiology **139**: 652–663.

Bogs J, Jaffe FW, Takos AM, Walker AR, Robinson SP. 2007. The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. PlantPhysiology **143**: 1347–1361.

Bontpart T, Cheynier V, Ageorges A, Terrier N. 2015. BAHD or SCPL acyltransferase ? What a dilemma for acylation in the world of plant phenolic compounds. New Phytologist **208**: 695–707.

Boulekbache-Makhlouf L, Meudec E, Chibane M, Mazauric JP, Slimani S, Henry M, Cheynier V, Madani K. 2010. Analysis by high-performance liquid chromatography diode array detection mass spectrometry of phenolic compounds in fruit of *Eucalyptus globulus* cultivated in Algeria. Journal of Agricultural and Food Chemistry **58**: 12615-24.

Cadot Y, Miñana-Castelló MT, Chevalier M. 2006. Anatomical, histological, and histochemical changes in grape seeds from *Vitis vinifera* L. cv Cabernet franc during fruit development. Journal of Agricultural and Food Chemistry **54**: 9206–15.

Carrier G, Huang YF, Le Cunff L, Fournier-Level A, Vialet S, Souquet JM, Cheynier V, Terrier N, This P. 2013. Selection of candidate genes for grape proanthocyanidin pathway by an integrative approach. Plant Physiology and Biochemistry **72**: 87-95.

Cavallini E, Matus JT, Finezzo L, Zenoni S, Loyola R, Guzzo F, Schlechter R, Ageorges A, Arce-Johnson P, Tornielli GB. 2015. The phenylpropanoid pathway is controlled at different branches by a set of R2R3-MYB C2 repressors in grapevine. Plant Physiology **167**:1448–70.

Cheynier V. 2005. Polyphenols in foods are more complex than often thought. American Journal of Clinical Nutrition **81**: 223S–229S.

Coombe BG, Hale CR. 1973. The hormone content of ripening grape berries and the effects of growth substance treatments. Plant Physiology **51**: 629–634.

Da Silva PPA, Laranjinha JA, de Freitas VA. 2003. Antioxidant protection of low density lipoprotein by procyanidins: structure/activity relationships. Biochemical Pharmacology **66**: 947–954.

Debeaujon I, Léon-Kloosterziel KM, Koornneef M. 2000. Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. Plant Physiology **122**: 403–414.

Della-Cioppa G, Bauer SC, Klein BK, Shah DM, Fraley RT, Kishore GM. 1986. Translocation of the precursor of 5-enolpyruvylshikimate-3-phosphate synthase into chloroplasts of higher plants in vitro. Proceedings of the National Academy ofSciences, USA **83**: 6873–6877.

Dewick PM, Haslam E. 1969. Phenol biosynthesis in higher plants. The Biochemical Journal **113**:537–42.

Dewick PM. 2002. The shikimate pathway: aromatic amino acids and phenylpropanoid pathway. In**:** Medicinal Natural Products. Paul M Dewick, John Wiley & Sons, Ltd.

Ding L, Hofius D, Hajirezaei MR, Fernie AR, Börnke F, Sonnewald U. 2007. Functional analysis of the essential bifunctional tobacco enzyme 3-dehydroquinate dehydratase/shikimatedehydrogenase in transgenic tobacco plants.Journal of Experimental Botany **58:** 2053–2067.

Dixon RA, Xie DY, Sharma SB. 2005. Proanthocyanidins: a final frontier in flavonoid research? New Phytologist **165**: 9–28.

Downey MO, Harvey JS, Robinson SP. 2003. Analysis of tannins in seeds and skins of Shiraz grapes throughout berry development. Australian Journal of Grape and Wine Research **9**: 15–27.

Emanuelsson O, Nielsen H, von Heijne G. 2006. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Science **8**: 978–984.

Eudes A, Bozzo GG, Waller JC, Naponelli V, Lim EK, Bowles DJ, Gregory JF 3rd, Hanson AD. 2008. Metabolism of the folate precursor p-aminobenzoate in plants: glucose ester formation and vacuolar storage. The Journal of Biological Chemistry **283**:15451–9.

Fabre S, Pinaud N, Fouquet E, Pianet I. 2010. Colloidal behavior of wine galloylated tannins. Comptes Rendus Chimie **13:** 561–565.

Ganson RJ, D'Amato TA, Jensen RA. 1986. The two-isozymesystem of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthasein *Nicotiana silvestris* and other higher plants. Plant Physiology **82**: 203–210.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research **31**:3784–8.

Guo J, Carrington Y, Alber A, Ehlting J. 2014. Molecular characterization of quinate and shikimate metabolism in *Populus trichocarpa*. Journal of Biological Chemistry **289**:23846–58.

Hause B, Meyer K, Viitanen PV, Chapple C, Strack D. 2002. Immunolocalization of 1-*O*-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. Planta **215**:26–32.

Herrmann KM, Weaver LM. 1999. The shikimate pathway. Annual Review of Plant Physiology and Plant Molecular Biology **50**:473–503.

Huang YF, Doligez A, Fournier-Level A, Le Cunff L, Bertrand Y, Canaguier A, Morel C, Miralles V, Veran F, Souquet JM, Cheynier V, Terrier N, This P. 2012. Dissecting genetic architecture of grape proanthocyanidin composition through quantitative trait locus mapping. BMC Plant Biology **12**:30.

Huang YF, Vialet S, Guiraud JL, Torregrosa L, Bertrand Y, Cheynier V, This P, Terrier N. 2014. A negative MYB regulator of proanthocyanidin accumulation, identified through expression quantitative locus mapping in the grape berry. New Phytologist **201**: 795–809.

Ikegami A, Sato A, Yamada M, Kitajima A, Yonemori K. 2005. Expression of genes involved in proanthocyanidin biosynthesis during fruit development in a Chinese pollination-constant, nonastringent (PCNA) persimmon, 'Luo Tian Tian Shi'. Journal of American Society of Horticultural Science **130:** 830–835.

Ikegami A, Eguchi S, Kitajima A, Inoue K, Yonemori K. 2007. Identification of genes involved in proanthocyanidin biosynthesis of persimmon (*Diospyros kaki*) fruit. Plant Science **172**: 1037–1047.

Jiang X, Liu Y, Li W, Zhao L, Meng F, Wang Y, Tan H, Yang H, Wei C, Wan X, Gao L, Xia T. 2013. Tissue-specific, development-dependent phenolic compounds accumulation profile and gene expression pattern in tea plant [*Camellia sinensis*]. PLoS One **8**:e62315.

Karimi M, Inze D, Depicker A. 2002. GATEWAY((TM)) vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science **7:** 193–195.

Kennedy JA, Troup GJ, Pilbrow JR, Hutton D.R, Hewitt D, Hunter CR, Ristic R, Iland PG, Jones GP. 2000. Development of seed polyphenols in berries from *Vitis vinifera* L. cv. Shiraz. Australian Journal of Grape Wine Research **6**: 244-254.

Khater F, Fournand D, Vialet S, Meudec E, Cheynier V, Terrier N. 2012. Identification and functional characterization of cDNAs coding for hydroxybenzoate/hydroxycinnamate glucosyltransferases co-expressed with genes related to proanthocyanidin biosynthesis. Journal of Experimental Botany **63:** 1201–1214.

Koyama K, Numata M, Nakajima I, Goto-Yamamoto N, Matsumura H, Tanaka N. 2014. Functional characterization of a new grapevine MYB transcription factor and regulation of proanthocyanidin biosynthesis in grapes. Journal of Experimental Botany **65**: 4433–49.

Kruger MJ, Davies N, Myburgh KH, Lecour S. 2014. Proanthocyanidins, anthocyanins and cardiovascular diseases. Food Research International **59**: 41–52.

Lambert M, Meudec E, Verbaere A, Mazerolles G, Wirth J, Masson G, Cheynier V, Sommerer N. 2015. A High-Throughput UHPLC-QqQ-MS Method for Polyphenol Profiling in Rosé Wines. Molecules **20**: 7890–914.

Landete JM. 2011. Ellagitannins, ellagic acid and their derived metabolites: A review about source, metabolism, functions and health. Food Research International **44**: 1150–1160.

Lee VS, Dou J, Chen RJ, Lin RS, Lee MR, Tzen JT. 2008. Massive accumulation of gallic acid and unique occurrence of myricetin, quercetin, and kaempferol in preparing old oolong tea. Journal of Agricultural and Food Chemistry **56:**7950–6.

Li K, Mikola MR, Draths KM, Worden RM, Frost JW. 1999. Fedbatch fermentor synthesis of 3-dehydroshikimic acid using recombinant *Escherichia coli*. Biotechnology and Bioengineering **64:**61–73.

Liu Y, Gao L, Liu L, Yang Q, Lu Z, Nie Z, Wang Y, Xia T. 2012. Purification and characterization of a novel galloyltransferase involved in catechin galloylation in the tea plant (*Camellia sinensis*). The Journal of Biological Chemistry **287:** 44406-44417.

Lizarraga D, Lozano C, Briedé JJ, van Delft JH, Touriño S, Centelles JJ, Torres JL, Cascante M. 2007. The importance of polymerization and galloylation for the antiproliferative properties of procyanidin-rich natural extracts. FEBS Journal **274**: 4802–4811.

Mané C, Souquet JM, Ollé D, Verriés C, Véran F, Mazerolles G, Cheynier V, Fulcrand H. 2007. Optimization of simultaneous flavanol, phenolic acid, and anthocyanin extraction from grapes using an experimental design: application to the characterization of champagne grape varieties. Journal of Agricultural and Food Chemistry **55**:7224–33.

McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R. 2013. Analysis Tool Web Services from the EMBL-EBI. Nucleic acids research 41 (Web Server issue):W597-600.

Michel G, Roszak AW, Sauvé V, Maclean J, Matte A, Coggins JR, Cygler M, Lapthorn AJ. 2003. Structures of shikimate dehydrogenase AroE and its Paralog YdiB. A common structural framework for different activities. The Journal of Biological Chemistry **278**:19463-72.

Milo R, Last RL. 2012. Achieving diversity in the face of constraints: lessons from metabolism. Science **336**: 1663-1667.

Monteiro F, Sebastiana M, Pais MS, Figueiredo A. 2013. Reference gene selection and validation for the early responses to downy mildew infection in susceptible and resistant *Vitis vinifera* cultivars. PLoS One **8**:e72998.

Mousdale DM, Campbell MS, Coggins JR. 1987. Purification and characterization of a bifunctional dehydroquinase-shikimate: NADP oxidoreductase from pea seedlings. Phytochemistry **26**:2665–2670.

Mugford ST, Qi X, Bakht S, Hill L, Wegel E, Hughes RK, Papadopoulou K, Melton R, Philo M, Sainsbury F, Lomonossoff GP, Roy AD, Goss RJ and Osbourn A. 2009. A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. Plant Cell **21:** 2473–2484.

Mugford ST, Louveau T, Melton R, Qi X, Bakht S, Hill L,Tsurushima T, Honkanen S, Rosser SJ, Lomonossoff GP, Osbourn A. 2013. Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat. The Plant Cell **25**: 1078–1092.

Muir RM, Ibáñez AM, Uratsu SL, Ingham ES, Leslie CA, McGranahan GH, Batra N, Goyal S, Joseph J, Jemmis ED, Dandekar AM. 2011. Mechanism of gallic acid biosynthesis in bacteria (*Escherichia coli*) and walnut (*Juglans regia*). Plant Molecular Biology **75**:555–65.

Niemetz R, Gross GG. 2005. Enzymology of gallotannin and ellagitannin biosynthesis. Phytochemistry **66**: 2001-11.

Nishizaki Y, Yasunaga M, Okamoto E, Okamoto M, Hirose Y, Yamaguchi M, Ozeki Y, Sasaki N. 2013. p-Hydroxybenzoyl-glucose is a zwitter donor for the biosynthesis of 7-polyacylated anthocyanin in Delphinium. Plant Cell **25**:4150–4165.

Ossipov V, Salminen JP, Ossipova S, Haukioja E, Pihlaja K. 2003. Gallic acid and hydrolysable tannins are formed in birch leaves from an intermediate compoundof the shikimate pathway. Biochemical Systematics and Ecology **31**: 3–16.

Pang Y, Cheng X, Huhman DV, Ma J, Peel GJ, Yonekura-Sakakibara K, Saito K, Shen G, Sumner LW, Tang Y, Wen J, Yun J, Dixon RA. 2013. Medicago glucosyltransferase UGT72L1: potential rolesin proanthocyanidin biosynthesis. Planta **238**:139-154.

Richman JE, Chang YC,Kambourakis S, Draths KM, Almy E, Snell KD, Strasburg GM, Frost JW. 1996. Reaction of 3-dehydroshikimic acid with molecular oxygen and hydrogen peroxide: products, mechanism, and associated antioxidant activity. Journal of American Chemical Society **118**:11587-11591.

Schmid J, Schaller A, Leibinger U, Boll W, Amrhein N. 1992. The in vitro synthesized tomato shikimate kinase precursor is enzymatically active and is imported and processed to the matureenzyme by chloroplasts. The Plant Journal **2**: 375–383.

Serrano M, Wang B, Aryal B, Garcion C, Abou-Mansour E, Heck S, Geisler M, Mauch F, Nawrath C, Métraux JP. 2013. Export of salicylic acid from the chloroplast requires the multidrug and toxin extrusion-like transporter EDS5. Plant Physiology **162**: 1815–1821.

Singh SA, Christendat D. 2006. Structure of Arabidopsis dehydroquinate dehydratase-shikimate dehydrogenase and implications for metabolic channeling in the shikimate pathway. Biochemistry **45**:7787-96.

Singh SA, Christendat D. 2007. The DHQ-dehydroshikimate-SDH-shikimate-NADP(H) complex: insights into metabolite transfer in the shikimate pathway. Crystal Growth & Design **7**:2153–2160.

Singh S, Stavrinides J, Christendat D, Guttman DS. 2008. A phylogenomic analysis of the shikimate dehydrogenases reveals broadscale functional diversification and identifies one functionally distinct subclass. Molecular Biology and Evolution **25**:2221–2232.

Song CP, Guo Y, Qiu Q, Lambert G, Galbraith DW, Jagendorf A, Zhu JK. 2004. A probable $Na^+(K^+)/H^+$ exchanger on the chloroplast envelope functions in pH homeostasis and chloroplast development in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences **101**:10211–6.

Souquet JM, Cheynier V, Brossaud F, Moutounet M. 1996. Polymeric proanthocyanidins from grape skins. Phytochemistry **43**: 509-512.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution **30**: 2725-2729.

Taneyama M. 1992. Studies on C-Glycosides in higher plants intracellular distribution of bergenin and related compounds in *Saxifraga stolonifera* leaves.The botanical magazine, Tokyo **105**:565–571.

Terrier N, Glissant D, Grimplet J, Barrieu F, Abbal P, Couture C, Ageorges A, Atanassova R, Léon C, Renaudin JP, Dédaldéchamp F, Romieu C, Delrot S, Hamdi S. 2005. Isogene specific oligo arrays reveal multifaceted changes in gene expression during grape berry (*Vitis vinifera* L.) development. Planta **222**:832-47.

Terrier N, Torregrosa L, Ageorges A, Vialet S, Verriès C, Cheynier V, Romieu C. 2009. Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. Plant Physiology **149**: 1028–1041.

Torregrosa L, Bouquet A. 1997. Agrobacterium rhizogenes and A. tumefaciens co-transformation to obtain grapevine hairy roots producing the coat protein of grapevine chrome mosaic nepovirus. Plant Cell, Tissue and Organ Culture **49**: 53-62.

Verriès C, Guiraud JL, Souquet JM, Vialet S, Terrier N, Ollé D. 2008. Validation of an extraction method on whole pericarp of grape berry (*Vitis vinifera* L. cv Syrah) to study biochemical and molecular aspects of flavan-3-ol synthesis during berry development. Journal of Agricultural Food Chemistry **56**: 5896–5904.

Vidal S, Cartalade D, Souquet J, Fulcrand H, Cheynier V. 2002. Changes in proanthocyanidin chain-length in wine-like model solutions. Journal of Agricultural Food Chemistry **50**: 2261–2266

Werner I, Bacher A, Eisenreich W. 1997. Retrobiosynthetic NMR studies with 13C-labeled glucose. Formation of gallic acid in plants and fungi. Journal of Biological Chemistry **272**:25474-82.

Werner RA, Rossmann A, Schwarz C, Bacher A, Schmidt HL, Eisenreich W. 2004. Biosynthesis of gallic acid in *Rhus typhina*: discrimination between alternative pathways from natural oxygen isotope abundance. Phytochemistry **65**:2809-13.

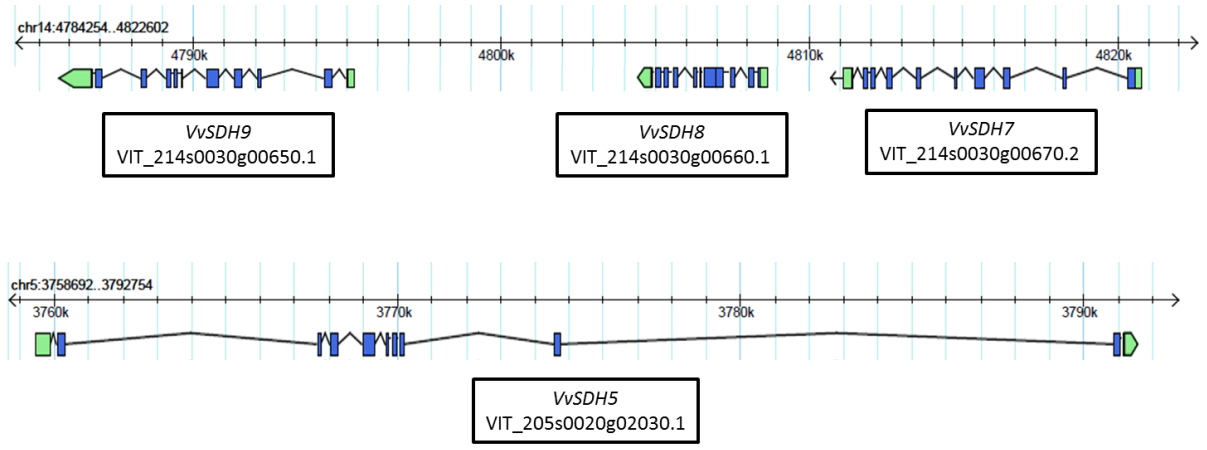
Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiology **126**: 485–493.

R Development Core Team. 2009. R: a language and environment for statistical computing, Vienna, Austria: R Foundation for Statistical Computing.

Yilmaz Y, Toledo RT. 2004. Major flavonoids in grape seeds and skins: antioxidant capacity of catechin, epicatechin, and gallic acid. Journal of Agricultural and Food Chemistry **52**: 255–60.

Supplemental Figure 1. Genomic structure of *VvSDHs*.

The exon-intron pattern was recovered from the 12X version of *Vitis vinifera* genome with gene prediction V2 (http://genomes.cribi.unipd.it/grape/). The scale indicates the chromosome and the genomic region shown.
Light green and blue boxes indicate the putative UTR regions and exonic sequences, respectively.



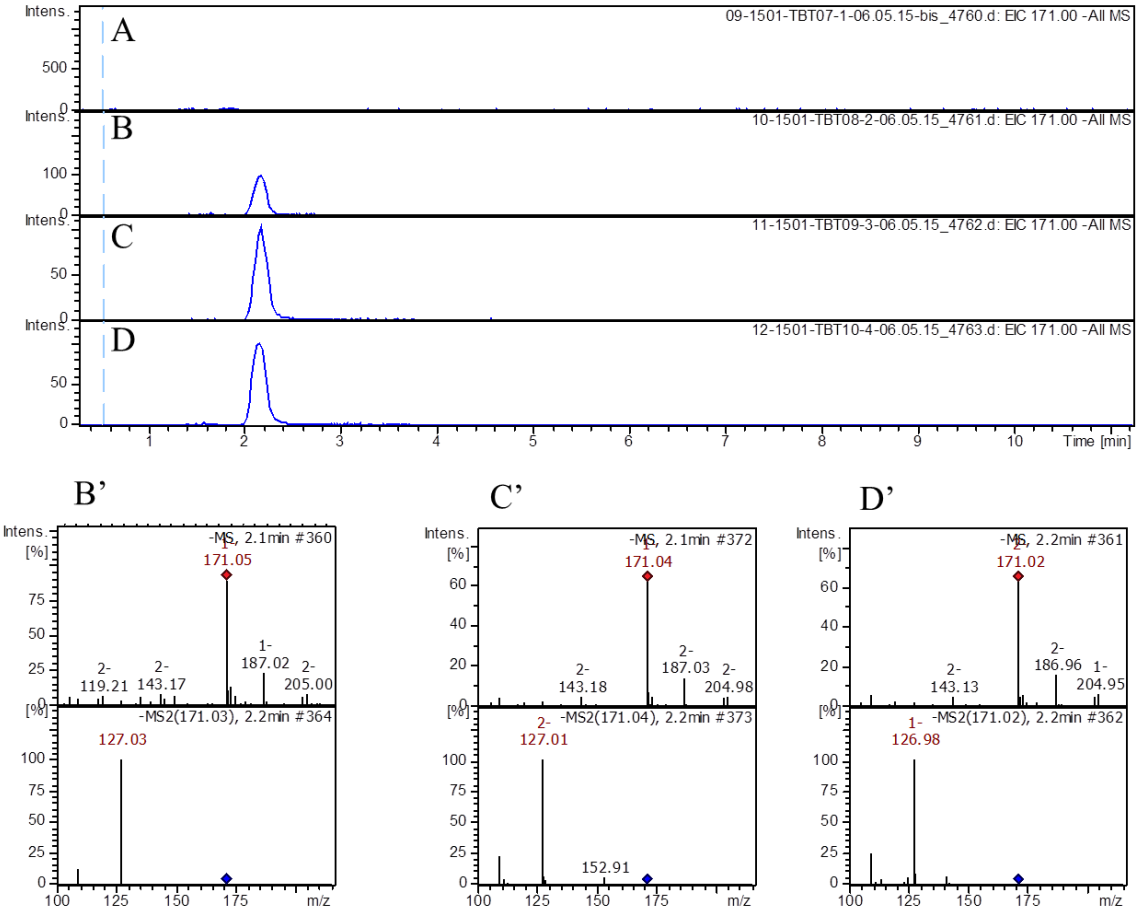
Supplemental Figure 2. Sequence identity of grape shikimate dehydrogenases (VvSDHs).

Amino acids and nucleotides (between brackets) identity is expressed in percentage and has been defined using ClustalW2 with default parameters (http://www.ebi.ac.uk/Tools/msa/clustalw2/) from cDNA sequences cloned from Syrah cultivar.

	VvSDH5	VvSDH7	VvSDH8	VvSDH9
VvSDH5	100	51.45 (59.17)	69.75 (72.44)	70.33 (74.68)
VvSDH7		100	46.82 (58.14)	48.94 (57.95)
VvSDH8			100	70.62 (74.62)
VvSDH9				100

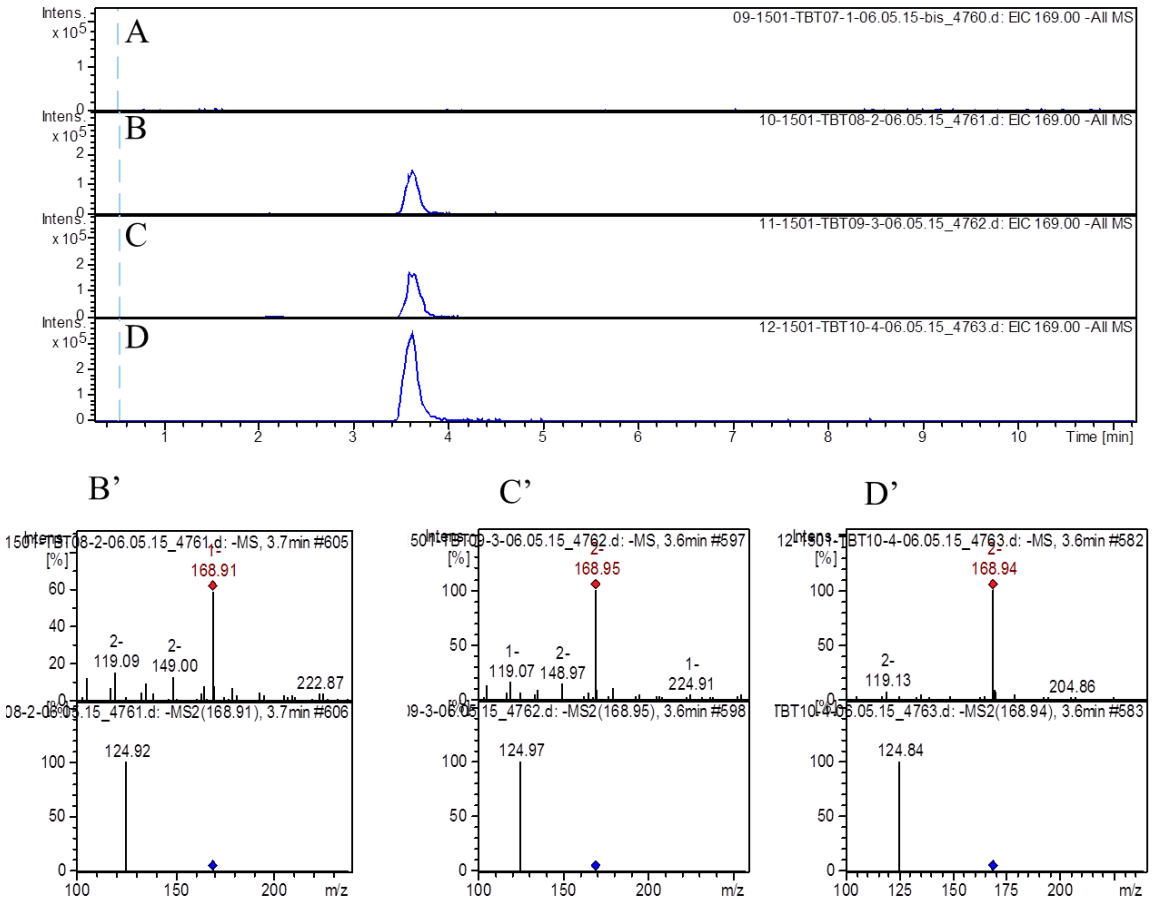
Supplemental Figure 3. UPLC-DAD-MS analysis of 3-dehydroshikimate in enzymatic assays.

Extracted-ion chromatogram at m/z 171in the negative ion modein enzymatic assay without enzyme from (A) shikimic acid and $NADP^+$, (B) 3-dehydroshikimate and $NADP^+$; and in presence of VvSDH8 from (C) shikimic acid and $NADP^+$, (D) 3-dehydroshikimate and $NADP^+$. The corresponding mass spectra and fragmentation by MS2 of 171 ion yielding a fragment ion at m/z127 is indicated by the same letter with an apostrophe.

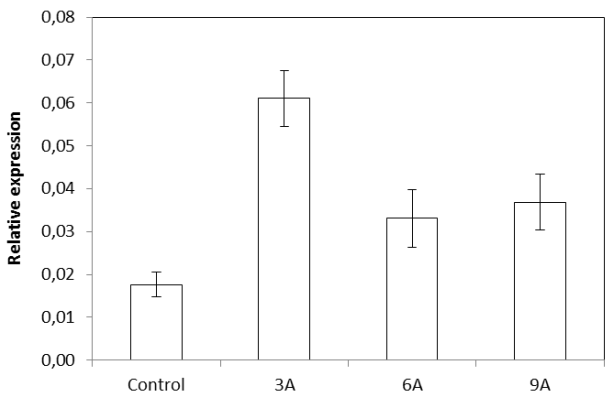


Supplemental Figure 4. UPLC-DAD-MS analysis of gallic acid in enzymatic assays.

Extracted-ion chromatogramat m/z 169 in the negative ion mode in enzymatic assay without enzyme from (**A**) shikimic acid and $NADP^+$, (**B**) 3-dehydroshikimate and $NADP^+$; and in presence of VvSDH8 from (**C**) shikimic acid and $NADP^+$, (**D**) 3-dehydroshikimate and $NADP^+$. The corresponding mass spectra and fragmentation by MS2 of 169 ion yielding a fragment ion at m/z125 is indicated by the same letter with an apostrophe.



Supplemental Figure 5. *VvSDH8* relative expression level in hairy-roots.
VvSDH8 expression was determined by real-time PCR and normalized with the expression of the reference gene *Elongation factor 1 α* (*EF1α*). Control line is devoid of transgene. The lines 3A, 6A and 9A are three independant transgenic lines transformed with the construct pH2GW7-VvSDH8. Data represent the mean of three replicates. Error bars indicate ± standard deviation.

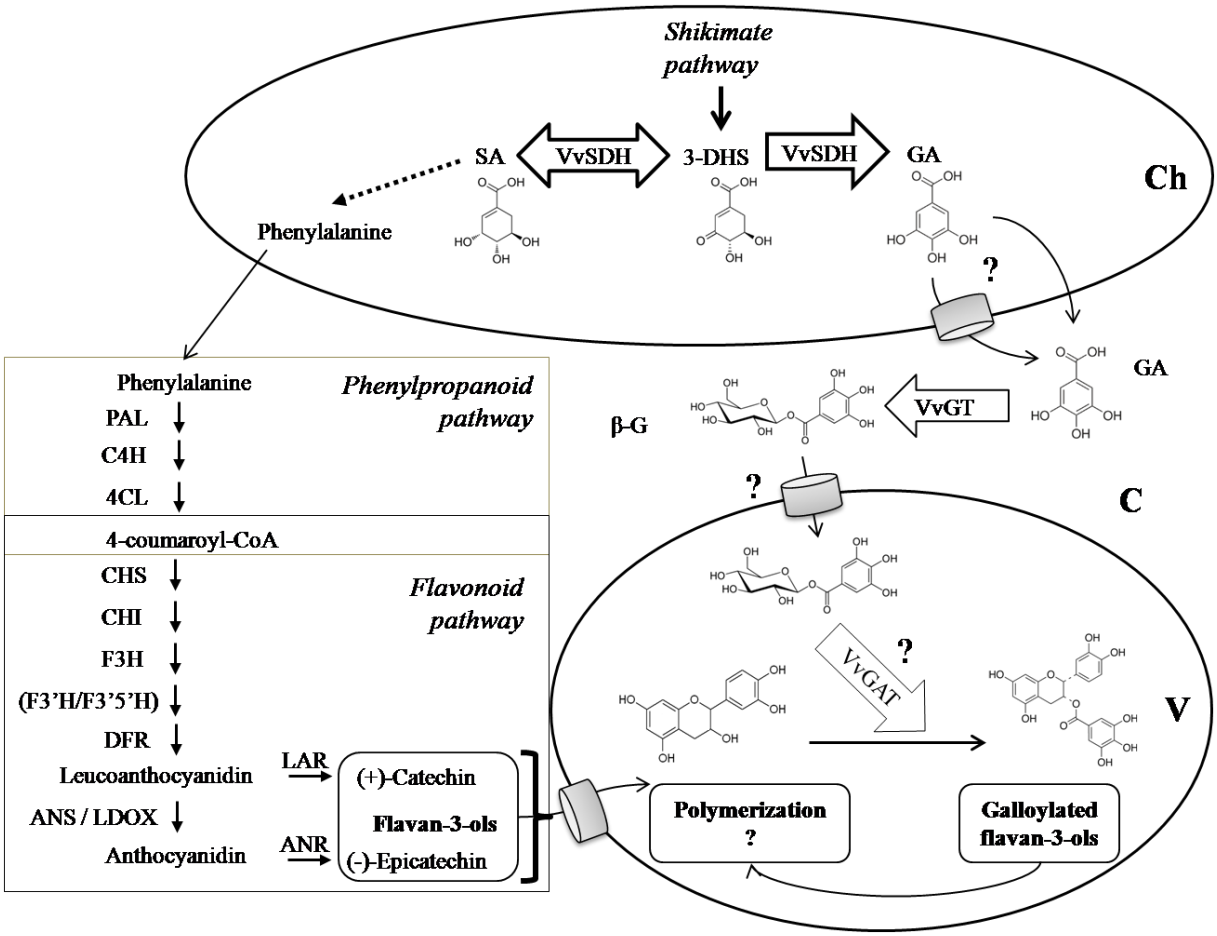


Supplemental Figure 6. Two-step galloylation of flavan-3-ols in grape.

3-Dehydroshikimate (3-DHS), a metabolite from the shikimate pathway, can be converted by *Vitis vinifera* shikimate dehydrogenases (VvSDHs) in the chloroplastic compartment (Ch) to provide shikimic acid (SA). The final steps of the shikimate pathway convert SA into phenylalanine which is a precursor for the phenylpropanoid pathway. The 4-coumaroyl-CoA produced is used in the flavonoid pathway to produce different flavonoids. At the end of the flavonoid pathway, (+)-catechin is synthesized from leucoanthocyanidin by the leucoanthocyanidin reductase (LAR) and (-)-epicatechin is synthesized from anthocyanidin by the anthocyanidin reductase (ANR). Flavan-3-ols (or uncharacterized precursors) are transported in the vacuole (V) where they polymerize *via* an unknown mechanism to form proanthocyanidins.

3-DHS can also be converted by VvSDHs into gallic acid (GA) in the chloroplast. Being a weak acid, GA could freely cross the chloroplast membranes or be routed by a chloroplastic transporter towards the cytosolic compartment (C). GA is glycosylated by specific grape Glucosyltranferases (VvGTs) leading to β-glucogalline (β-G). This glucose ester of GA could be transported in the vacuole by a tonoplastic transporter. Galloyl moiety from β-glucogalline could be used by a grape Glucose Acyltransferase (VvGAT) to acylate flavan-3-ols monomers in the vacuolar compartment, producing galloylated flavan-3-ols.

Other abbreviations are: PAL, phenylalanine ammonia-lyase; C4H, cinnamate-4-hydroxylase; 4CL, 4-coumarate CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3'H, flavonoid-3'-hydroxylase; DFR, dihydroflavonol reductase; ANS, Anthocyanidin synthase; LDOX, leucoanthocyanidin dioxygenase.



Supplemental Table 1. Primers used for *SDH* cloning and quantitative polymerase chain reaction (qPCR).

Purpose	Name	5'-3' sequence	Sense	Restriction site
pGEMT	SDH5pGEXstart	GGATCC**ATG**GAAAGCGGAGGAATGAG	F	BamH
Easy cloning	SDH5pGEXstop	CTCGAGTTATTGCAAATTTGATATGAATTGCTT	R	XhoI
	SDH7pGEXstart	GGATCC**ATG**GATGATGTTGGAGTTTTGAA	F	BamHI
	SDH7pGEXstop	CTCGAGTCAGAACTTTGATAAAATAATCTCC	R	XhoI
	SDH8pGEXstart	GGATCC**ATG**GGGAGCCTCCCATTTACTGT	F	BamHI
	SDH8pGEXstop	CTCGAGTTATGCGTGTTTCGACATAAGCTC	R	XhoI
	SDH9pGEXstart	GGATCCGGAGCCCGGAGAAATTCG	F	BamHI
	SDH9pGEXstop	CTCGAGTTATGTATTCCTCACCAAAACTTCC	R	XhoI
pH2GW7 cloning	SDH8TopoStart	CACC**ATG**GGGAGCCTCCCATTTACTGT	F	
	SDH8TopoStop	TTATGCGTGTTTCGACATAAGCTCC	R	
	SDH9TopoStart	CACC**ATG**ACTCTCAGCAGCGTTCCG	F	
	SDH9TopoStop	TTATGTATTCCTCACCAAAACTTCC	R	
Quantitative PCR	SDH5Q	GGCAGGCATATGAGCAGTTT	F	
Grape berry	SDH5Q	CAGCTCTCAACAGGACAGCTC	R	
	SDH7Q	AGTTCTGATACACTGCCTCATG	F	
	SDH7Q	ACCCTTATTGAAATGTGCTTGGA	R	
	SDH8Q	AGAGATGTTGATCCGCCAAG	F	
	SDH8Q	AACTGACTGCCCAAGCAAAT	R	
	SDH9Q	TCAGGGAAGTTTTGGTGAGGA	F	
	SDH9Q	ACTCAGCAGCTAATCACAAAGA	R	
Quantitative PCR	SDH8OEQ	AGTTGCCAACCGAACATTTGA	F	
Hairy-roots	SDH8OEQ	GCTTGGGAATGGGAGTGTCA	R	

Start codon is in bold letters.

Supplemental Table 2. List of quantified compounds and MRM parameters (ion mode, precursor and product ions m/z, retention times).

	Ion Mode	m/z precursor	m/z quantifier	m/z qualifier	retention time (min)
Direct injection					
Gallic acid	-	168.9	125.0	79.0	1.80
β-glucogallin	-	331.0	169.0	151.0	1.78
Phloroglucinolysis					
Gallocatechin	+	307.1	139.0	151.0	3.40
Epigallocatechin	+	307.1	139.0	151.0	4.93
(Epi)gallocatechin phloro	+	431.2	127.0	305.1	2.08
Catechin	+	291.1	139.0	123.1	4.96
Catechin phloro	+	415.1	127.1	289.1	3.56
Epicatechin	+	291.1	139.0	123.1	5.48
Epicatechin phloro	+	415.1	127.1	289.1	4.06
Epicatechin gallate	+	443.1	123.1	273.1	5.78
Epicatechin gallate phloro	+	567.2	153.1	247.2	5.01

Abbreviations: Phloro: phloroglucinol adduct.

Supplemental Table 3. Summary of plant species used to analyze DQD/SDH sequences.

Species	Common name	Family	ID	Accession n°
Arabidopsis thaliana	Arabidopsis	Brassicaceae	AtSDH	AAF08579
Camellia sinensis	Tea plant	Theaceae	CasSDH1	AIZ93902
			CasSDH2	AJA40947
			CasSDH3	AJA40948
Diospyros kaki	Persimmon	Ebenaceae	DkSDH1	BAI40147
Eucalyptus grandis	Eucalyptus	Myrtaceae	EgSDH1	Eucgr.H01214.1
			EgSDH2	Eucgr.H04428.1
			EgSDH3	Eucgr.H04427.1
			EgSDH4	Eucgr.B01770.2
			EgSDH5	Eucgr.J00263.6
Fragaria vesca	Strawberry	Rosaceae	FvSDH1	XP_004302480
			FvSDH2	XP_004302479
			FvSDH3	XP_004289250
			FvSDH4	XP_004288087
Juglans regia	Walnut	Juglandaceae	JrSDH	AAW65140
Nicotiana tabacum	Tobacco	Solanaceae	NtSDH1	AAS90325
			NtSDH2	AAS90324
Populus trichocarpa	Poplar	Salicaceae	Poptr1	Potri.010G019000
			Poptr2	Potri.013G029900
			Poptr3	Potri.005G043400
			Poptr4	Potri.014G135500
			Poptr5	Potri.013G029800
Solanum lycopersicum	Tomato	Solanaceae	SlSDH1	AAC17991
			SlSDH2	XP_010327280
			SlSDH3	XP_004242317
Citrus sinensis	Orange	Rutaceae	CsSDH1	orange1.1g010050m
			CsSDH2	orange1.1g010101m
			CsSDH3	orange1.1g007151m
Vitis vinifera	Grapevine	Vitaceae	VvSDH5	VIT_205s0020g02030
			VvSDH7	VIT_214s0030g00670
			VvSDH8	VIT_214s0030g00660
			VvSDH9	VIT_214s0030g00650

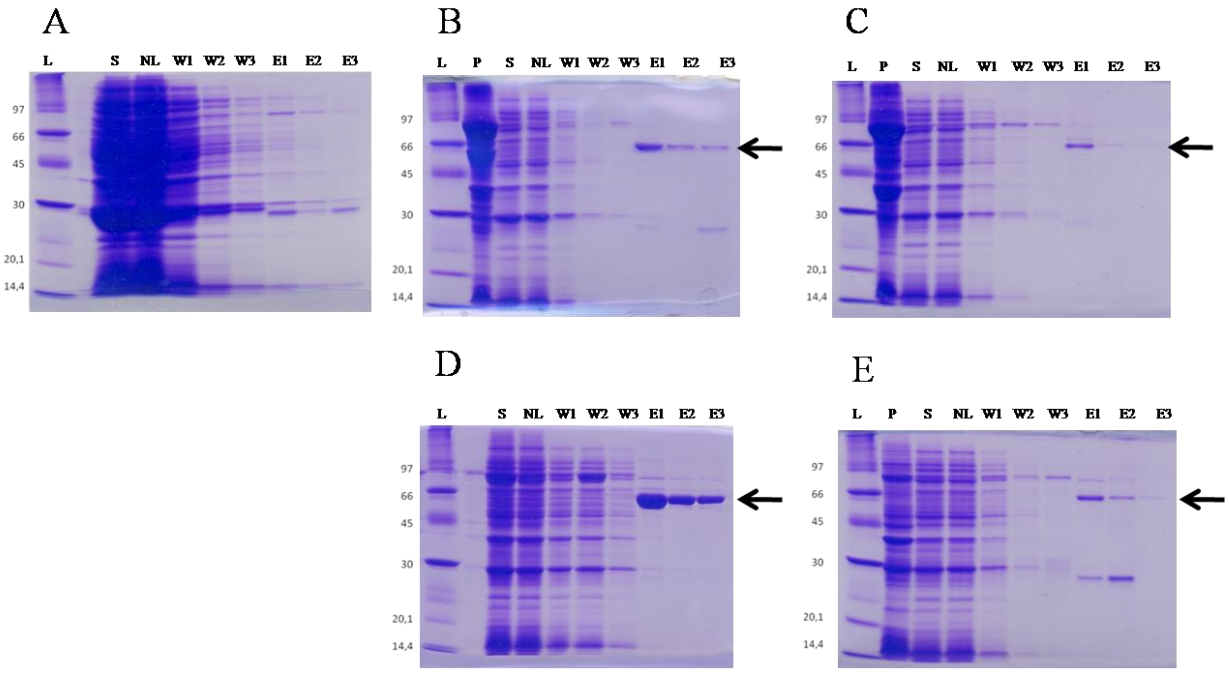


Figure 26. Analysis of recombinant VvSDHs on SDS-PAGE gel.
The ladder shows the molecular weight in kDa.A. Empty vector pGEX-4T2. B. VvSDH5. C. VvSDH7. D. VvSDH8. E. VvSDH9.
Abbreviations: L: ladder, P: pellet, S: supernatant, NL: Non-linked, W1-3: Wash 1-3, E1-3: Eluate 1-3.

3. Other results

3. 1. SDS-PAGE analysis

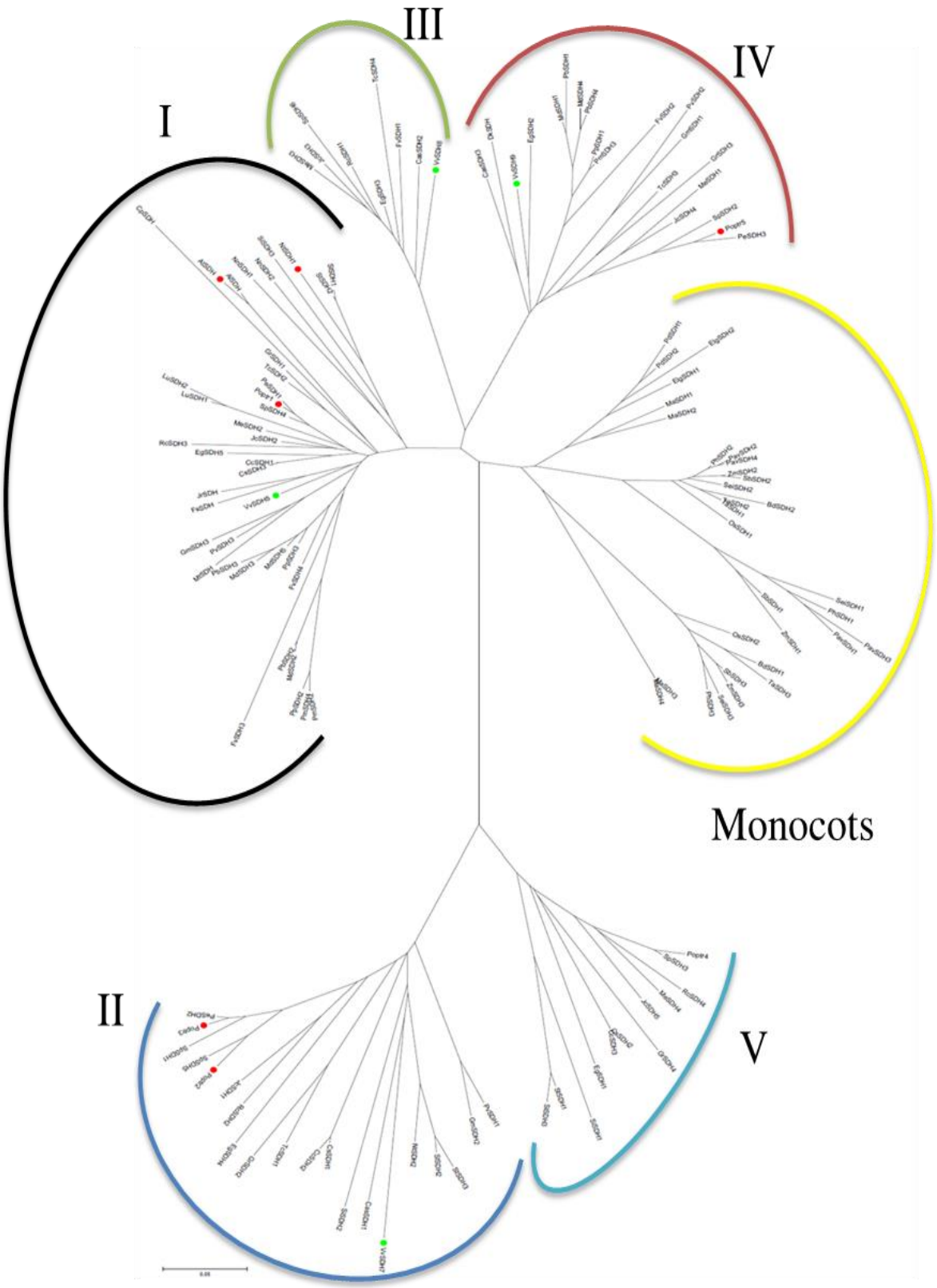
The different fractions leading to VvSDHs production and purification described in Material and methods have been analysed on SDS-PAGE gel to check the purity of the fractions and the molecular weight of the purified proteins before enzymatic assay (Fig. 26).The fractions E1, E2 and E3 correspond to the purified purified proteins. No band corresponding to the expected molecular weight of the cloned VvSDH sequences (~60 kDa) was observed with protein extracted from bacteria transformed with empty vector (Fig. 26, A). This band (designated with an arrow) was observed for each vector in which one of the 4 studied VvSDH was cloned (Fig. 26, B-E).

3. 2. Phylogenetic analysis of shikimate dehydrogenases in plants

To extend the phylogenetic analysis shown in manuscript #1, we collected other DQD/SDH sequences from plants from public databases.
Hence, a second neighbor-joining tree has been constructed from 100 sequences from 34 Dicots and 31 sequences from 11 Monocots (Fig. 27, A). Sequence alignement and prediction of putative signal peptide are presented in Supplemental Figure 1 and 5, respectively (see Annexes).
The distribution of DQD/SDH genes is summarised in Table 13.
Additional dicot sequences clustered in one of the 5 groups already described in manuscript #1, suggesting that the divergences among sequences are common to all dicot species.

All Monocots species from which DQD/SDH sequences have been collected belong to Poaceae family, except banana (*Musa accuminata*, Musaceae), African oil palm (*Elaeis guineensis*, Arecaceae) and date palm (*Phoenix dactylifera*, Arecaceae).
These sequences clustered in the "Monocot group", independently of Dicot sequences. Monocots possess between 1 and 4 DQD/SDH genes. 4 distinct groups have been identified among the "Monocot group": Monocots (M)-I, M-II, M-III and M-IV (Fig. 27, B).
Arecacea and Musaceae DQD/SDHs clustered in group M-I, except MaSDH3 and -4 which grouped independently of these 4 groups. Sorghum (*Sorghum bicolor*), maize (*Zea mays*), *Panicum hallii* and *Setaria italica* possess one gene in groups M-II, M-III and M-IV.
Brachypodium distachyon, rice (*Oryza sativa*) and wheat (*Triticum aestivum*) possess genes in groups M-II and M-IV.
Panicum virgatum possess 2 genes in clades M-II and M-III.

A



3. 3. Comparison of key amino acids

The key amino acids involved in substrate binding and catalysis, and cofactor binding have been studied for Arabidopsis SDH (Singh & Christendat, 2006). The 31 Monocot DQD/SDH sequences were aligned with AtSDH to compare key catalytic and substrate-binding motifs (Fig. 28). Contrary to Dicot sequences, few divergences at key amino acids were observed. The divergences were found in SxS motif and Thr 381. PavSDH3 is devoid of Thr 381 and presents a Gln instead of Lys 385. The members of group M-IV exhibit a Ser instead of Thr 381.
MaSDH3 and -4, not included in one of the 4 groups, exhibit a Thr instead of Ser 336.
In group M-I, ElgSDH1 and -2 exhibit a Gly and an Arg instead of Ser 338, respectively. In PhdSDH1 sequence, Thr 381 is replaced by a Gly as observed in Groups II and V Dicots sequences.
The 3 other key motifs were very well conserved in Monocots SDHs sequences.

4. Discussion

GA formation in plants

The SA pathway routes the carbon flux towards aromatic amino acid biosynthesis, crucial for plant development. DQD/SDH is the only bifunctional enzyme of this pathway in which it ensures the 3rd and the 4th enzymatic steps. Due to the fusion of DQD and SDH domains, this enzyme must take kinetics advantage of the proximity of these two domains (Moore, 2004). Hence, SDH activity was reported 9-fold higher than DQD activity (Hermann, 1995), that is why the vicinity of the two domains efficiently routes the carbon flux towards phenylalanine. In other words, 3-DHS synthesized by the DQD domain can be directly used by the SDH domain to produce SA, and avoid the loss of carbon structures.

However, the SA pathway can be subject to enzymatic branchpoints leading to the biosynthesis of (hydroxy)benzoic acids (reviewed by Widhalm & Dudareva, 2015). The knowledge concerning putative genes involved in GA biosynthesis from SA pathway metabolites is relatively limited because a single study reported the capacity of a SDH to produce GA from SA and 3-DHS (Muir *et al*., 2011).

Here, we show that the 4 genes encoding DQD/SDH in grapevine genome exhibit different catalytic properties from SA and 3-DHS.
VvSDH5 exhibits a “classical” SDH activity, converting efficiently 3-DHS to SA. “Classical” SDH activity must minimize the spontaneous formation of GA and PCA from 3-DHS. Indeed, we observed a relative instability of 3-DHS leading tothe spontaneous formation of these HBAs in alkaline conditions. This phenomenon could occur *in planta* because the SA pathway is confined in chloroplasts where pH is rather alkaline. Otherwise, this conversion can be catalysed by ions such as Cu^{2+}, as demonstrated *in vitro* (Kambourakis & Frost, 2000).
Otherwise, the “classical” SDH activity can lead to a small production of GA, as demonstrated from SA and $NADP^+$.We don’t know whether it results from the spontaneous conversion of the enzymically formed 3-DHS into GA or if VvSDH5 catalyses this conversion.
Two genes (VvSDH8 and -9) exhibited lower “classical” SDH activity than VvSDH5 but catalysed GA production from 3-DHS and $NADP^+$ *in vitro*. To sum up, GA formation could result from enzymatic and non-enzymatic process *in planta*.

B

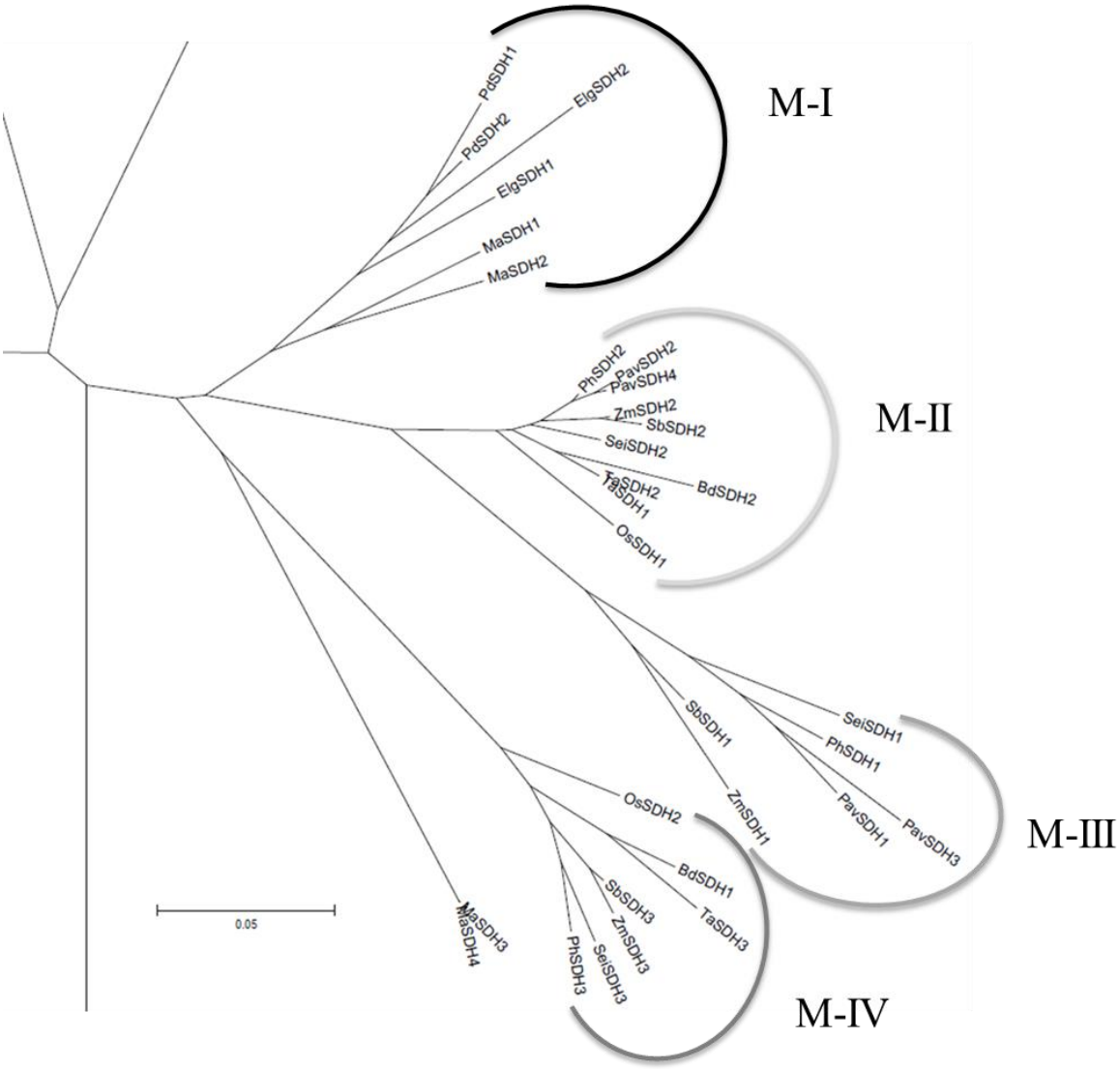


Figure 27. Shikimate dehydrogenase neighbor-joining tree.
Red circle: characterized DQD/SDH, green circles: VvSDHs. **A**. Dicots and Monocots, **B**. Monocots. This phylogenetic tree was constructed from the 4 VvSDHs sequenced in this study and 127 sequences available on public databases (NCBI and Phytozome) using MEGA6 software. The scale bar represents 0.05 substitutions per site.

The intracellular localization of VvSDHs in transformed grapevine hairy-roots could be further experimented. Indeed, the lack of chloroplastic transit peptide in some VvSDH sequences could lead to SDH activity in cytosol rather than in chloroplasts. This hypothesis involves the export of substrates (3-dehydroquinate, 3-DHS, SA) from chloroplast to cytosol. If those metabolites can freely cross plastic envelops or if transporters are involved would require further investigations.

Common genetic bases for GA biosynthesis in plants?

Few species carry a single *DQD/SDH* gene (e.g. *Arabidopsis* sp.). AtSDH exhibits a "classical" SDH activity (Singh & Christendat, 2006) which avoids the formation of GA *in planta*. No study reported the presence of GA in *Arabidopsis thaliana* and this metabolite has never been detected in this plant to date (Isabelle Debeaujon, personnal communication). Hence, this metabolite is not or very little accumulated in this plant, attesting for a carbon flux efficiently canalysed in the SA pathway.
Nevertheless, several *DQD/SDH* genes (between 1 and 6) were found in dicot genomes. This multiplication of *DQD/SDH* in plant genomes could reflect the genes evolution by duplication as shown for several genes involved in the SA pathway in woody species (Tohge *et al*., 2013).
Side by side localization of 3 VvSDHs in chromosome 14 must result from tandem array duplication, a mechanism frequently observed in plant genomes (Flagel & Wendel, 2009). Duplicated genes of the SA pathway could acquire a new function as proposed for poplar (Hamberger *et al*., 2006). The recent characterization of 4 poplar DQD/SDHs argue in favour of a substrate (SA or quinic acid) and/or cofactor ($NADP^+$ or NAD^+) specificity of these enzymes. Hence, depending on the catalytic properties of DQD/SDH, SA pathway metabolites could be re-oriented for the biosynthesis of secondary metabolites, including GA.
The present phylogenetic study reveals that dicot DQD/SDHs can be classified in 5 major groups.
Hence, we could predict the role played by the members of the same group from the measured activity of VvSDH *in vitro*. We observed common divergences in key amino acids for SDH activity among the members of the 5 groups, which illustrates the specification of DQD/SDH for substrate and cofactor. The enzymes from groups I, III and IV are predicted to have higher affinity for SA and $NADP^+$.

Group I genes were found in all completely sequenced genomes, that attests the importance of those genes, due to their hight "classical" SDH activity. Unexpectedly, the GA-producer walnut SDH clustered in group I instead of groups III or IV. Nevertheless, the not entirely sequenced walnut genome could carry other DQD/SDHs clustering in other groups.

Group III genes are rare because only 9 species possess a gene from this group. No other gene from this group was previously characterized. In contrary, 18 species (53%) carry at least one gene from group IV. Interestingly, some plants which produce galloylated flavan-3-ols possess one gene from group IV (e.g. tea plant, persimmon). Poplar Poptr5 was the only characterized group IV enzyme. Unfortunately, GA production was not quantified for poplar Poptr5 characterization (Guo *et al*., 2014). Muir and coworkers (2011) also tested GA production after incubating protein extracts from different plant species in presence of SA and $NADP^+$. Interestingly, the species carrying at least one gene in group III or IV (cotton plant, apple tree) produced higher GA content than other plant species (*Medicago sativa*, tobacco, tomato, rice).
Due to their lower "classical" SDH activity and their capacity to produce GA, enzymes in groups III and IV must take part of GA metabolism. The apparition of new DQD/SDH genes in some dicot genomes must be involved in the accumulation of GA and GA-derived metabolites (galloylated tannins).

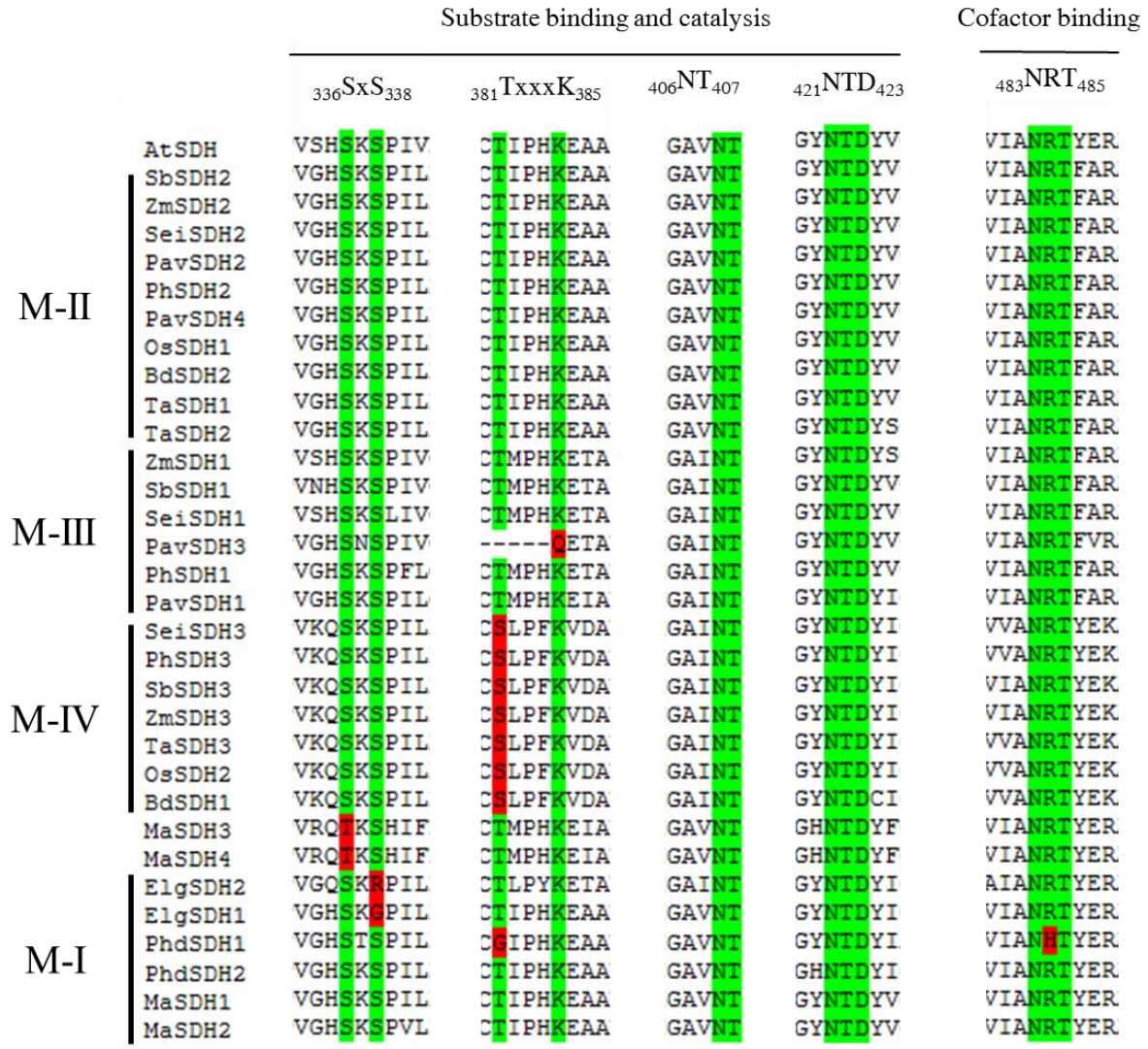


Figure 28. Alignment of key catalytic and substrate-binding motifs in the SDH domain of Monocots DQD/SDH.
Protein alignement has been performed by ClustalW. Amino acids numerotation is based on DQD/SDH from *Arabidopsis thaliana* (AtSDH, Singh & Christendat, 2006). Key amino acids are shaded in green when identical to AtSDH, and in red when different from AtSDH. The groups M-I to M-IV have been determined by neighbor-joining tree.

Between 1 and 4 genes encoding DQD/SDH have been collected from monocot genomes. However, key motifs involved in substrate binding and orientation are better conserved than in Dicot SDHs. The substitution of Thr381 by a Ser in the members of group M-IV could affect substrate affinity.
Contrary to Dicots, ellagic acid and hydrolysable (GA based) tannins are absent or very rare in Monocots (Bate-Smith, 1968, 1972, 1973).The analysis of Monocots DQD/SDH sequences and comparison with Dicots was a pertinent approach to support the hypotheses made from Dicots.
Key motifs among SDH domain are similar to those of Arabidopsis SDH, which could explain why GA and GA-based metabolites are few or not biosynthesized in Monocots.

CHAPTER 4: GLUCOSE ACYLTRANSFERASES

1. Introduction

Flavonoid biosynthesis is recognized to occur at the cytosolic side of the endoplasmic reticulum by an enzymatic complex (Winckel-Shirley, 2001). Once synthesized, these molecules can be subject to various substitutions in their carbon structure: hydroxylation, methylation, glycosylation, acylation. It was suggested that these "decorations" are used to target molecules to their storage compartment, avoiding the dispersion of flavonoids within the cell. Indeed, due to their high reactivity, these molecules could be toxic and upset the cellular machinery. Otherwise, those metabolites must be sufficiently accumulated to ensure their biological role *in planta*. The affinity of vesicular or tonoplastic transporter for "tagged" (i.e. acylated) flavonoids is consistent with this hypothesis (Gomez *et al*., 2009).
Acylation consists in the transfer of an acyl group (-COOH) of a "donor" molecule to an "acceptor" molecule. The donor molecule must be activated by an energy-rich link. Thus, acylation usually occurs from acyl-Coenzyme A thioesters or 1-*O*-β-glucose esters.
The enzymes involved in such reactions belong to two different families. Acylation from acyl-Coenzyme A thioesters is catalysed by BAHD acyltransferases (BAHD-ATs) whereas acylation from 1-*O*-β-glucose esters is catalysed by Serine Carboxypeptidases-Like acyltransferases (SCPL-ATs).

Comparison between BAHD and SCPL-ATs involved in the biosynthesis of phenolic compounds has been reviewed in a manuscript accepted in New Phytologist (Manuscript #2).

This chapter focusses on the study of 2 genes annotated as Serine Carboxypeptidases, named Glucose Acyltransferase 1 and -2 (VvGAT1 and -2). Transcriptomic analysis (Terrier *et al*., 2009) and quantitative genetic study (Huang *et al*., 2012) revealed that those 2 genes could be involved in PA biosynthesis.

These 2 VvGATs have been tested for their capacity to transfer a galloyl moiety on flavan-3-ols (galloylation) and hydroxycinnamoyl moieties on tartaric acid.

2. Manuscript # 2

New Phytologist

Review

Research review

BAHD or SCPL acyltransferase? What a dilemma for acylation in the world of plant phenolic compounds

Author for correspondence:
Nancy Terrier
Tel: +33 4 99 61 24 60
Email: terrier@supagro.inra.fr

Received: *10 March 2015*
Accepted: *6 May 2015*

Thibaut Bontpart, Véronique Cheynier, Agnès Ageorges and Nancy Terrier

INRA, UMR1083 SPO, 2, place, Viala, F-34060 Montpellier, France

New Phytologist (2015) **208:** 695–707
doi: 10.1111/nph.13498

Key words: acylation, acyltransferase, BAHD, biosynthesis, phenolic compounds, secondary metabolite, Serine CarboxyPeptidase-Like (SCPL).

Summary

Phenolic compounds are secondary metabolites involved in several plant growth and development processes, including resistance to biotic and abiotic stresses. The biosynthetic pathways leading to the vast diversity of plant phenolic products often include an acylation step, with phenolic compounds being the donor or acceptor molecules. To date, two acyltransferase families using phenolic compounds as acceptor or donor molecules have been described, with each using a different 'energy-rich' acyl donor. BAHD-acyltransferases, named after the first four biochemically characterized enzymes of the group, use acyl-CoA thioesters as donor molecules, whereas SCPL (Serine CarboxyPeptidase Like)-acyltransferases use 1-*O*-β-glucose esters. Here, common and divergent specifications found in the literature for both enzyme families were analyzed to answer the following questions. Are both acyltransferases involved in the synthesis of the same molecule (or same group of molecules)? Are both acyltransferases recruited in the same plant? How does the subcellular localization of these enzymes impact metabolite trafficking in plant cells?

Introduction

Plant phenolic compounds are ubiquitous secondary metabolites and include molecules with a phenolic ring, that is, a benzyl ring with at least one hydroxyl group. These phenolic compounds include hydroxybenzoic acids (HBAs), hydroxycinnamic acids (HCAs) and more complex polyphenols, for example, flavonoids. Among them, flavonoids are one of the most important subgroups, with *c.* 6000 molecules with multiple substitutions, including glycosylation and acylation (Harbone & Williams, 2000).

In planta, phenolic compounds protect against biotic and/or abiotic stresses. In addition, these compounds are involved in the coloration of seeds and flowers, which attracts different pollinators or disseminators (Cheynier *et al.*, 2013). Acylation can influence the function of phenolic compounds *in planta*. Notably, acylation impacts different properties of anthocyanins, such as solubilization, stabilization, vacuolar uptake and color modulation (Nakayama *et al.*, 2003). Malate acylation with a sinapoyl moiety confers anti-UV properties (Landry *et al.*, 1995). In oat, acylation of avenacins impacts their involvement in plant defense (Mugford *et al.*, 2009). Acylation can also modulate the transport and storage of molecules (Gomez *et al.*, 2009; Zhao & Dixon, 2009a).

Chemically, acylation is the transfer of an acyl group via an activated donor molecule to an acceptor molecule. This reaction is endergonic, thus requiring an 'energy-rich' acyl donor. Acylation extends to several classes of secondary compounds and hence expands metabolite diversity, involving molecules from various pathways (Mugford & Milkowski, 2012). Interestingly, phenolic compounds can be donor or acceptor molecules. Those identified as donor molecules are phenolic acids, such as HCAs, which have a C6-C3 skeleton (e.g. *p*-coumaric acid), or HBAs, which have a C6-C1 skeleton (e.g. gallic acid). Flavonoids, among others, are acceptor molecules and thus can be subject to acylation.

Two enzyme families using phenolic compounds govern the acylation of secondary compounds in vascular plants: BAHD-acyltransferases (BAHD-ATs), named after the first four biochemically characterized enzymes of the family, and Serine CarboxyPeptidase-Like (SCPL)-acyltransferases (SCPL-ATs). BAHD-ATs use acyl-CoA thioesters as 'energy-rich' donor molecules, whereas SCPL-ATs use 1-*O*-β-glucose esters (Fig. 1). In

2006, D'Auria classified genetically and biochemically characterized BAHD-ATs. To complement this review, we have extended the list of enzymes to include BAHD-ATs that are involved in the acylation of phenolic compounds and to include SCPL-ATs genetically and biochemically characterized. We exclude those BAHD-ATs using benzoates to acylate non-phenolic compounds, as reported in Schmidt *et al.* (2015). Common and divergent specifications of these two AT families are described and illustrated using different studies.

BAHD-ATs

Specifications

The BAHD-AT family takes its name from the first biochemically characterized enzymes, benzylalcohol *O*-acetyltransferase (**B**EAT) (Dudareva *et al.*, 1998), anthocyanin *O*-hydroxycinnamoyltransferase (**A**HCT) (Fujiwara *et al.*, 1998), anthranilate *N*-hydroxycinnamoyl/benzoyltransferase (**H**CBT) (Yang *et al.*, 1997) and deacetylvindoline 4-*O*-acetyltransferase (**D**AT) (St-Pierre *et al.*, 1998). Varying from 48 to 55 kDa, BAHD-ATs function as monomers. These enzymes share two conserved motifs: an HxxxD motif involved in substrate binding and the less conserved DFGWG motif, which probably plays a structural rather than a catalytic role (Unno *et al.*, 2007). Despite these motifs, members of the BAHD-AT family share a low protein sequence identity (10–30%) (St-Pierre & De Luca, 2000). In 2006, D'Auria classified 46 BAHD-ATs into five major clades. Generally, there is no donor molecule specificity within a single clade.

More recently, a phylogenetic analysis of 69 biochemically characterized BAHD-ATs and putative BAHD-ATs of five model plants suggested the existence of eight clades instead of five (Tuominen *et al.*, 2011). This study reveals that a plant genome contains numerous genes coding for BAHD-ATs (from 50 to 100 depending on the species) and highlights new inter- and intra-clade motifs.

Our focus on BAHD-ATs that use phenolic compounds resulted in the counting of new characterized enzymes in addition to those described by D'Auria (2006). We divided these enzymes into six major groups.

Characterized BAHD-ATs

Flavonoid BAHD-ATs Flavonoids share a common C6-C3-C6 carbon structure. *In planta*, most flavonoids are glucosides. In 2006, the 12 BAHD-ATs able to acylate flavonoids were classified in clade I, with the exception of Ss5MaT2 (D'Auria, 2006). These BAHD-ATs share the NYFGNC diagnostic motif (Nakayama *et al.*, 2003). We have inventoried 12 new flavonoid BAHD-ATs, most of which acylate anthocyanins, although some acylate other flavonoids (Table 1).

Characterized aromatic ATs transfer hydroxycinnamoyl moieties, whereas known aliphatic ATs transfer the malonyl moiety. They both acylate the glucosyl residue of flavonoid glycosides but exhibit specificity related to: (1) the position in which the glucose is attached to the flavonoid; (2) the position of the hydroxyl on the glucose; and (3) the type of acyl-CoA, as illustrated in the following examples.

Most identified enzymes are aliphatic ATs, such as Ss5MaT1 from *Salvia splendens* (Suzuki *et al.*, 2001). To date, 19 flavonoid malonyltransferases have been characterized (Table 1). Ss5MaT1 malonylates bisdemalonylsalvianin and shisonin by transferring a malonyl moiety to the hydroxyl in the C6 position of the 5-*O*-glucosyl residue to produce monodemalonylsalvianin and malonylshisonin, respectively (Fig. 2). Ss5MaT2 also acylates the 5-*O*-glucosyl residue, but on the hydroxyl group in the C4 position instead of the C6. A malonyltransferase identified in tobacco (NtMAT1) acylates both flavones and flavonol glucosides (Taguchi *et al.*, 2005).

Four aromatic ATs have been identified: Gt5AT (Fujiwara *et al.*, 1998), Pf3AT (Yonekura-Sakakibara *et al.*, 2000), At3At1 and At3At2 (Luo *et al.*, 2007). Gt5AT is able to transfer caffeoyl or *p*-coumaroyl residues to the C6 hydroxyl of the 5-*O*-glucosyl residue of anthocyanin 3,5-diglucosides. At3At1 and At3At2 exhibit affinity for *p*-coumaroyl-, feruloyl- and caffeoyl-, but not

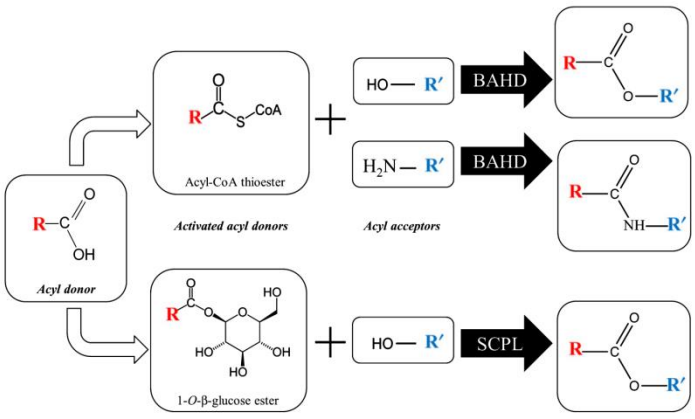


Fig. 1 Acylation mechanisms catalyzed by BAHD and SCPL acyltransferases. The carboxylic group (COOH) of the acyl donor (R) can be esterified with Coenzyme A (SCoA) or glucose to form an activated acyl donor. Acyl-CoA thioesters and 1-O-β-glucose esters are 'energy-rich' molecules used by BAHD and SCPL acyltransferases, respectively. Both types of enzyme can acylate the hydroxyl group (OH) of the acyl acceptor (R′) (*O*-acylation). In some cases, BAHD acyltransferases can also acylate the amine group (NH_2) of the acyl acceptor (*N*-acylation).

Table 1 (Continued)

Name	Acyl-CoA donor	Major formed product(s)	Species	Reference(s)	Clade
CbHST	Hydroxycinnamoyl	Hydroxycinnamoyl shikimates, hydroxybenzoates and hydroxyanthranilate	*Coleus blumei* Benth	Sander & Petersen (2011)	nd
CbRAS	Caffeoyl or *p*-coumaroyl	Hydroxycinnamoyl hydroxyphenyllactates	*Coleus blumei* Benth	Berger *et al.* (2006)	nd
MoRAS	Caffeoyl or *p*-coumaroyl	Hydroxycinnamoyl hydroxyphenyllactates	*Melissa officinalis*	Weitzel & Petersen (2011)	nd
LaAT1	Hydroxycinnamoyl	Hydroxycinnamoyl acid amides, hydroxycinnamoyl shikimates and 4-hydroxyphenyllactate	*Lavandula angustifolia*	Landmann *et al.* (2011)	V
Alkyl-hydroxycinnamate ester BAHD acyltransferases					
AtHHT/ASFT	Feruloyl	Suberin aromatics	*Arabidopsis thaliana*	Gou *et al.* (2009); Molina *et al.* (2009)	Va
FHT	Feruloyl	Suberin aromatics	*Solanum tuberosum*	Serra *et al.* (2010)	nd
PtFHT1	Feruloyl, *p*-coumaroyl	Alkyl hydroxycinnamates	*Populus trichocarpa*	Cheng *et al.* (2013)	nd
DCF	Feruloyl, sinapoyl	ω-Hydroxy fatty acids (cutin)	*Arabidopsis thaliana*	Rautengarten *et al.* (2012)	Va
FACT	Caffeoyl	Alkyl hydroxycinnamates	*Arabidopsis thaliana*	Kosma *et al.* (2012)	Va
Cell wall BAHD acyltransferases					
OsPMT	*p*-Coumaroyl	Monolignols, hydroxycinnamates	*Oryza sativa*	Withers *et al.* (2012)	Va
OsAT10	*p*-Coumaroyl	Glucuronoarabinoxylan (GAX)	*Oryza sativa*	Bartley *et al.* (2013)	Va
BdPMT	*p*-Coumaroyl	Monolignol *p*-coumarate conjugates	*Brachypodium distachyon*	Petrik *et al.* (2014)	V

*Already identified by D'Auria (2006); nd, non-determined. Abbreviations and GenBank accession numbers are listed below.

Flavonoid BAHD acyltranferases: Gt5AT, *Gentiana triflora* anthocyanin 5-aromatic acyltransferase (BAA74428); Dv3MAT, *Dahlia variabilis* malonyl-CoA:anthocyanidin 3-*O*-glucoside-6″-*O*-malonyltransferase (AAO12206); Sc3MaT, *Senecio cruentus* malonyl-CoA:anthocyanidin 3-*O*-glucoside 6″-*O*-malonyltransferase (AAO38058); Dm3MAT1, *Dendranthema* × *morifolium* anthocyanidin 3-*O*-glucoside-6″-*O*-malonyltransferase 1 (AAQ63615); Dm3MAT2 (AAQ63616); Dm3MAT3 (BAF50706); NtMAT1, *Nicotiana tabacum* malonyl-CoA flavonoid/naphthol glucoside acyltransferase (BAD93691); Vh3MAT1, *Verbena* × *hybrida* malonyl-CoA:flavonol 3-*O*-glucoside-6″-*O*-malonyltransferase (AAS77403); Lp3MAT1, *Lamium purpureum* malonyl-CoA:flavonol 3-*O*-glucoside-6″-*O*-malonyltransferase (AAS77404); Pf3AT, *Perilla frutescens* hydroxycinnamoyl-CoA:anthocyanin 3-*O*-glucoside-6″-*O*-acyltransferase (BAA93475); Ss5MaT1, *Salvia splendens* malonyl-CoA:anthocyanin 5-*O*-glucoside-6″-*O*-malonyltransferase 1 (AAL50566); Ss5MaT2, *Salvia splendens* anthocyanin 5-*O*-glucoside-4″-*O*-malonyltransferase (AAR26385); Pf5MaT, *Perilla frutescens* malonyl-CoA:anthocyanin 5-*O*-glucoside-6″-*O*-malonyltransferase (AAL50565); At5Mat, *Arabidopsis thaliana* malonyl-CoA:anthocyanidin 5-*O*-glucoside-6″-*O*-malonyltransferase (NP_189600); At3At1, *Arabidopsis thaliana* hydroxycinnamoyl-CoA:anthocyanin 3-*O*-glucoside-6″-*O*-acyltransferase 1 (NP_171890); At3At2 (NP_171849); GmIF7MaT, *Glycine max* malonyl-CoA:isoflavone 7-*O*-glucoside-6″-*O*-malonyltransferase (BAF73620); GmMT7, *Glycine max* malonyltransferase (ABY59019); MtMat1, *Medicago truncatula* malonyl-CoA:flavonol 3-*O*-glucoside-6″-*O*-malonyltransferase 1 (ABY91220); MtMat2 (ABY91222); MtMat3 (ABY91221); MaT4 (ADV04046); MaT5 (ADV04047); MaT6 (ADV04048); AtPMat1, *Arabidopsis thaliana* phenolic glucoside malonyltransferase 1 (AAK96528); AtPMat2 (NP_189609).

Acyl-CoA *N*-acyltransferases: HCBT, anthranilate *N*-hydroxycinnamoyl/benzoyltransferase (CAB06430); AsHHT, *Avena sativa* hydroxycinnamoyl-CoA:hydroxyanthranilate *N*-hydroxycinnamoyltransferase (BAC78633); AtSDT, *Arabidopsis thaliana* spermidine disinapoyl transferase (NP_179932); AtSCT, *Arabidopsis thaliana* spermidine dicoumaroyl transferase (NP_180087); SHT, spermidine hydroxycinnamoyl transferase (NP_179497); ACT, agmatine coumaroyltransferase (AAO73071); AtACT, *Arabidopsis thaliana* agmatine coumaroyltransferase (At5g61160).

BAHD benzoyltransferases: CbBEBT, *Clarkia breweri* benzoyl-coenzyme A (CoA):benzyl alcohol benzoyl transferase (AAN09796); NtBEBT, *Nicotiana tabacum* benzoyl-coenzyme A (CoA):benzyl alcohol benzoyl transferase (AAN09798); PhBPBT, *Petunia* × *hybrida* benzoyl-CoA:benzyl alcohol/phenylethanol benzoyltransferase (AAU06226); PtSABT, *Populus trichocarpa* benzoyl-CoA:salicyl alcohol *O*-benzoyltransferase (AJD25144); PtBEBT, *Populus trichocarpa* benzoyl-CoA:benzyl alcohol *O*-benzoyltransferase (AJD25145).

Alkyl-hydroxycinnamate ester BAHD acyltransferases: AtHHT, *Arabidopsis thaliana* hydroxycinnamoyl-CoA:ω-hydroxyacid *O*-hydroxycinnamoyltransferase (ACY78659); ASFT, aliphatic suberin feruloyl transferase (AAL34170); FHT, fatty ω-hydroxyacid/fatty alcohol hydroxycinnamoyltransferase (ACS70946); PtFHT1, *Populus trichocarpa* hydroxyacid/fatty alcohol hydroxycinnamoyltransferase 1 (JX515962); DCF, Deficient in Cutin Ferulate (At3g48720); FACT, fatty alcohol:caffeoyl-CoA caffeoyl transferase (At5g63560).

HCT/HQT and relatives: NtHQT, *Nicotiana tabacum* hydroxycinnamoyl-CoA:quinate hydroxycinnamoyltransferase (CAE46932); NtHCT, *Nicotiana tabacum* hydroxycinnamoyl-CoA:shikimate/quinate hydroxycinnamoyltransferase (CAD47830); AtHCT, *Arabidopsis thaliana* hydroxycinnamoyl-CoA:shikimate/quinate hydroxycinnamoyltransferase (NP_199704); PrHCT, *Pinus radiata* hydroxycinnamoyl-CoA:shikimate/quinate hydroxycinnamoyltransferase (ABO52899); CcHQT, *Coffea canephora* hydroxycinnamoyl-CoA:quinate hydroxycinnamoyltransferase (ABO77957); CcsHQT, *Cynara cardunculus* var. *scolymus* hydroxycinnamoyl-CoA:quinate hydroxycinnamoyltransferase (ABK79689); HQT1, hydroxycinnamoyl-CoA:quinate hydroxycinnamoyltransferase 1 (CAM84302); HQT2 (ACJ23164); LjHQT, *Lonicera japonica* hydroxycinnamoyl-CoA:quinate hydroxycinnamoyltransferase (ACZ52698); PvHCT1a, *Panicum virgatum* hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase 1a (BAO20883); PvHCT2a (AGM90558); PvHCT-Like1 (AFY17066); TpHCT1A, *Trifolium pratense* hydroxycinnamoyl-CoA:shikimate/quinate hydroxycinnamoyltransferase 1 A (ACI16630); TpHCT1B (ACI28534); TpHCT2 (ACI16631); CsHCT, *Cucumis sativus* hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (AEJ88365); OsHCT4, *Oryza sativa* hydroxycinnamoyltransferase 4; SlHQT, *Solanum lycopersicum* hydroxycinnamoyl-CoA:quinate transferase (NP_001234850); PtHCT1 and 6: *Populus trichocarpa* hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase 1 and 6; CbHST, *Coleus blumei* Benth. hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase (CBI83579); CbRAS, *Coleus blumei* Benth. rosmarinic acid synthase also called hydroxycinnamoyl-CoA:hydroxyphenyllactate hydroxycinnamoyltransferase (CAK55166); MoRAS, *Melissa officinalis* rosmarinic acid synthase (CBW35684); LaAT1, *Lavandula angustifolia* acyltransferase 1 (ABI48360).

Cell wall BAHD acyltransferases: OsPMT, *Oryza sativa p*-coumarate monolignol transferase (Os01g18744); OsAT10, *Oryza sativa* acyltransferase 10 (Os06g39390); BdPMT, *Brachypodium distachyon p*-coumaroyl-CoA:monolignol transferase (Bradi2g36910).

Suberin is deposited on the inner side of the primary cell wall of secondary shoots and roots. A transferase named AtHHT (hydroxycinnamoyl-CoA:ω-hydroxyacid *O*-hydroxycinnamoyltransferase; Gou *et al.*, 2009), and ASFT (aliphatic suberin feruloyl transferase) by another team (Molina *et al.*, 2009), specifically acylates 16-hydroxypalmitic acid with feruloyl-CoA. Orthologous genes encoding fatty ω-hydroxyacid/fatty alcohol hydroxycinnamoyltransferase have been identified (Table 1).

In roots, lipophilic compounds associated with suberin are called root waxes and also contain hydroxycinnamates. Fatty alcohol: caffeoyl-CoA caffeoyl transferase takes part in the biosynthesis of root waxes in *A. thaliana* (Kosma *et al.*, 2012). Like DCF, these enzymes have been classified in clade Va (Table 1).

Cell wall BAHD-ATs The plant primary cell wall, composed, *inter alia,* of cellulose and hemicellulose, ensures both rigidity and plasticity of the cell, allowing growth and division. The secondary cell wall, stiffer, is further composed of lignin, a complex aromatic polymer. Members of Poaceae have a primary cell wall mainly composed of cellulose microfibrils sealed in a (glucurono)arabinoxylan matrix acylated by ferulic and *p*-coumaric acid. OsAt10, belonging to clade Va, has been characterized in rice as a *p*-coumaroyl-CoA transferase (Bartley *et al.*, 2013).

A branch of the phenylpropanoid pathway provides the monolignols necessary for the formation of lignin units, notably via HCT activity (cf. the Section on HCT/HQT and relatives). In grasses, HCAs can also be esterified with monolignols by BAHD-ATs (Withers *et al.*, 2012).

To conclude, BAHD-ATs classified in different clades (Tuominen *et al.*, 2011) can use the same acyl-CoA thioester (e.g. Gt5AT from clade Ib and AtHCT from clade Vb use hydroxycinnamoyl-CoA). However, each clade exhibits more specificity for the acyl acceptor (e.g. clade Ib has specificity for flavonoids), even if BAHD-ATs from the same clade can acylate a large array of secondary metabolites (e.g. clade Vb).

Intracellular localization

Several members of the BAHD family have been shown to be cytosolic (Fujiwara *et al.*, 1998; Yu *et al.*, 2008; Panikashvili *et al.*, 2009). In plants, acyl-CoA thioesters are synthesized during metabolic processes localized in different compartments (Shockey & Browse, 2011), but can be exported to the cytosol via transport mechanisms to generate a cytosolic pool (Agrimi *et al.*, 2012). Hence, the cytosolic pool of acyl-CoA is compatible with BAHD-AT activity in this compartment.

Most flavonoids are known to be stored in the vacuolar compartment. Malonylation has been proposed to increase the stability of anthocyanins and their uptake and trapping into the vacuole (Heller & Forkmann, 1994; Suzuki *et al.*, 2002). GmMT7, an isoflavonoid malonyltransferase from soybean, was localized in the cytosol (Dhaubhadel *et al.*, 2008). In *Medicago truncatula*, malonylation of flavonoid glucosides can be catalyzed by several BAHD-ATs (Yu *et al.*, 2008; Zhao *et al.*, 2011). MtMaT1 was localized in both the cytoplasm and the nucleus (Yu *et al.*, 2008). This unexpected result is congruent with the earlier observation of flavonoid accumulation in the nuclear compartment (e.g. Feucht *et al.*, 2004). However, whether the nucleus is a site of biosynthesis or storage for flavonoids remains unclear. Malonylation has been shown to promote flavonoid vacuolar import via a tonoplastic multidrug and toxin extrusion transporter, MATE2 (Zhao *et al.*, 2011). In *Vitis vinifera*, MATE-type transporters called anthoMATE specifically mediate the transport of acylated anthocyanins into the vacuole (Gomez *et al.*, 2009). AnthoMATE proteins have been localized on the tonoplast and are associated with cytoplasmic anthocyanin-filled vesicles (Gomez *et al.*, 2011). Hence, acylation can impact anthocyanin transport and storage within the plant cell (Fig. 3).

SCPL-ATs

Discovery of SCPL-ATs

The transacylation reaction from 1-*O*-β-glucose esters was first described several decades ago and is involved in the biosynthesis of diverse phenolic compounds, such as CGA (Kojima & Uritani, 1972), sinapoylesters (Dahlbender & Strack, 1984) and gallotannins (Gross, 1983). Nevertheless, the first gene encoding an enzyme able to transacylate from 1-*O*-β-glucose ester was not identified until the early 2000s (Li & Steffens, 2000). Unexpectedly, these ATs are homologous to Serine Carboxy-Peptidases (SCPs) and were therefore named 'SCP-Like (SCPL)'-ATs. To date, only a small number of SCPL-ATs have been characterized, and our list extends these results to include incompletely characterized proteins and proteins that do not use phenolic compounds (Table 2).

Specifications

SCP and SCPL-ATs belong to the α/β hydrolases superfamily and share a common motif composed of three non-successive amino acids: serine (Ser), His and aspartic acid (Asp). This catalytic triad is a recurrent motif necessary for hydrolysis and other related functions, and could explain the recruitment of SCPL-ATs during the evolution of secondary metabolite acylation (Milkowski & Strack, 2004). Nonetheless, SCPL-ATs have lost their peptidase activity. Bioinformatic analysis revealed that 51 SCPL genes encoded in the *A. thaliana* genome clustered into several clades (Fraser *et al.*, 2005). The 19 SCPL genes in clade IA, including characterized sinapoyltransferases, were predicted to have SCPL-AT function *in planta*.

Bioinformatic analysis of the characterized SCPL-AT sequences revealed an N-terminal signal peptide, suggesting that the protein is processed within a secretory pathway towards the vacuole (Lehfeldt *et al.*, 2000; Mugford *et al.*, 2009). This hydrophobic signal peptide is predicted to be removed from the preprotein (Li & Steffens, 2000). Shirley & Chapple (2003) expressed a sinapoylglucose:choline sinapoyltransferase (SCT) in yeast and reported additional protein processing that results in two polypeptides, which suggested that SCT could function as a heterodimer. Immunoblot analyses of SCPL1 from oat seedling roots or expressed in tobacco leaves have also demonstrated the presence of

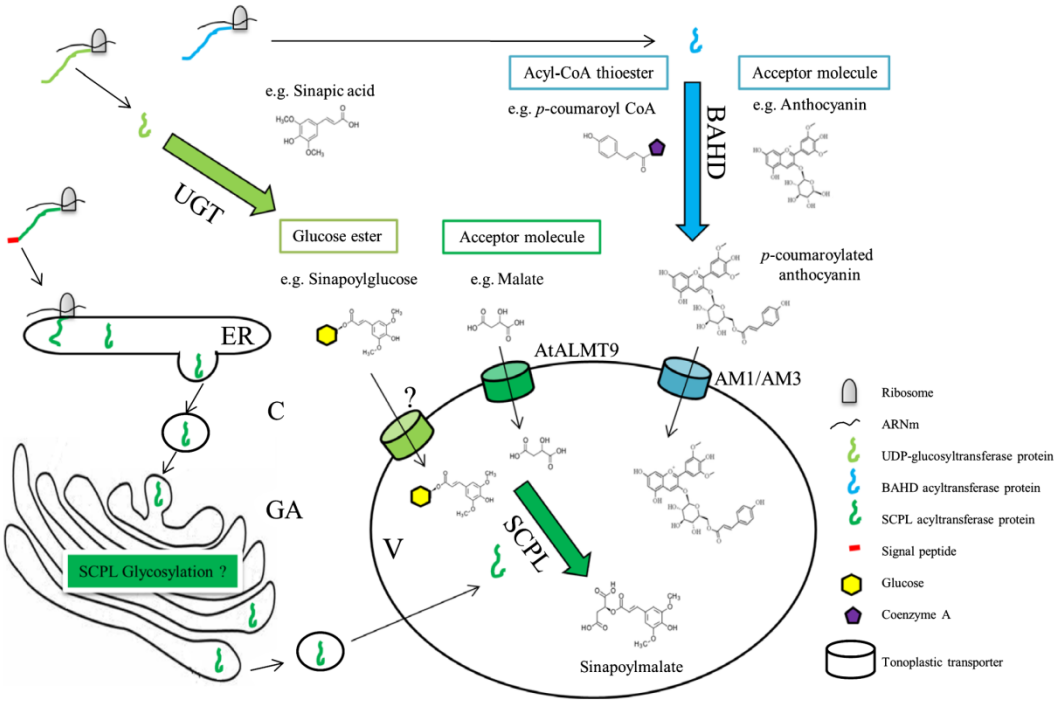


Fig. 3 Step-by-step BAHD and SCPL acylation mechanisms at the cellular level. BAHD acyltransferases use acyl-CoA (e.g. *p*-coumaroyl-CoA) to acylate acceptor molecules (e.g. anthocyanin) in the cytosolic compartment (C). The produced compound (e.g. *p*-coumaroylated anthocyanin) can be transported through the tonoplast via a specific transporter (e.g. antho multidrug and toxin extrusion transporter 1 (AM1) and 3 (AM3) from *Vitis vinifera*). UDP-Glucosyltransferase (UGT) uses UDP-glucose to glycosylate secondary metabolites and form glucose esters in the cytosolic compartment (e.g. sinapoyl glucose). Glucose esters can be imported into the vacuole (V) via tonoplastic transporters. In the vacuole, glucose esters are donor molecules used by SCPL acyltransferases to acylate acceptor molecules (e.g. malate). SCPL acyltransferases are predicted to contain a signal peptide that allows the protein to follow the secretory pathway before final targeting to the vacuole. Hause *et al.* (2002) inspired the route presented here for sinapoylglucose:malate sinapoyltransferase (SMT). The signal peptide is recognized and cleaved at the endoplasmic reticulum (ER). The protein is then trafficked to the Golgi apparatus (GA), where it probably undergoes glycosylation. Finally, the protein is targeted to the vacuole to catalyze acylation with glucose ester (e.g. malate sinapoylation). Malate can be imported into the vacuole by the aluminum-activated malate transporter from *Arabidopsis thaliana* (AtALMT9).

two polypeptides (Mugford *et al.*, 2009). SCPL1 sequence alignment with characterized SCP- and SCPL-ATs revealed the presence of an extended linker region in which cleavage potentially occurs. However, no post-translational cleavage has been noted for sinapoylglucose:malate sinapoyltransferase (SMT) from *A. thaliana* (Hause *et al.*, 2002). Nonetheless, a size discrepancy has been highlighted between SMT produced in *Escherichia coli* and in the plant. In fact, the SMT sequence carries six glycosylation sites (NXT/S) and is most likely glycosylated in the Golgi apparatus before being targeted to the vacuole *in planta* (Hause *et al.*, 2002; Fig. 3).

Because of the post-translational modifications described above and because eukaryotic hosts carry suitable cellular machinery for such modifications, the characterization of SCPL-ATs in eukaryotic hosts is needed. Heterologous expression of SCPL-ATs in yeast (Stehle *et al.*, 2008) and tobacco (Weier *et al.*, 2008; Mugford *et al.*, 2009) has proven to be a good strategy for obtaining functional SCPL-ATs.

Characterized SCPL-ATs

To date, only nine SCPL-ATs have been fully characterized (Table 2). All of these enzymes are from the Dicots, with the exception of SCPL1, which was identified in a monocotyledon (Mugford *et al.*, 2009). Like BAHD-ATs, SCPL-ATs interfere in diverse metabolic pathways to provide compounds with various biological roles. In *Lycopersicum penellii*, a glucose acyltransferase (GAC) catalyzes the formation of 1,2-di-*O*-acyl-β-glucose, a glucose isobutyryl polyester involved in insect resistance (Li & Steffens, 2000). In *A. thaliana*, four SCPL-ATs synthesize sinapate esters using 1-*O*-sinapoyl-β-D-glucose as donor molecule. SMT acylates malate from 1-*O*-sinapoyl-β-D-glucose to form sinapoylmalate (Lehfeldt *et al.*, 2000). Moreover, SCT catalyzes choline acylation from 1-*O*-sinapoyl-β-D-glucose to form sinapoylcholine (sinapine) in seeds (Shirley & Chapple, 2003). Two SCT orthologs (*BnSCT1* and *BnSCT2*) were identified in *Brassica napus* (Milkowski *et al.*, 2004; Weier *et al.*, 2008). 1-*O*-sinapoyl-β-

Table 2 Characterized and putative SCPL acyltransferases

Name	Glucose ester donor	Major formed product(s)	Species	Localization	Reference(s)
GAC	Isobutyryl	Glucose polyesters	*Solanum pennellii*		Li & Steffens (2000)*
CtAT1	*p*-Coumaroyl	Anthocyanin	*Clitoria ternatea*	Vacuole	Noda *et al.* (2007)
SMT	Sinapoyl	Sinapoylmalate	*Arabidopsis thaliana*	Vacuole	Lehfeldt *et al.* (2000); Hause *et al.* (2002)
SCT	Sinapoyl	Sinapoylcholine	*Arabidopsis thaliana*		Shirley & Chapple (2003)
BnSCT1	Sinapoyl	Sinapoylcholine	*Brassica napus*		Milkowski *et al.* (2004b)
BnSCT2	Sinapoyl	Sinapoylcholine	*Brassica napus*		Weier *et al.* (2008)
SAT	Sinapoyl	Sinapoylated anthocyanin	*Arabidopsis thaliana*		Fraser *et al.* (2007)
SST	Sinapoyl	1,2-Disinapoyl-glucose			
SCPL1 (SAD7)	*N*-Methyl anthraniloyl or benzoyl	Avenacin	*Avena strigosa*	Vacuole	Mugford *et al.* (2009, 2013)*
–	Hydroxycinnamoyl	Chlorogenic acid	*Ipomoea batatas*		Villegas & Kojima (1986)
–	Galloyl	Gallotannins	*Quercus* sp.		Gross (1983); Niemetz & Gross (2005)
–	Indole-3-acetyl	Indole acetyl-myo-inositol	*Zea mays*		Kowalczyk *et al.* (2003)*
–	Hydroxycinnamoyl	Anthocyanins	*Daucus carota*		Glaessgen & Seitz (1992)
AMalT	Malyl	Anthocyanins	*Dianthus caryophyllus*		Abe *et al.* (2008)
DkSCPL1	Galloyl	Galloylated flavan-3-ols	*Diospyros kaki* Thunb.		Ikegami *et al.* (2007)
DkSCPL2					Akagi *et al.* (2009)
VvGAT1	Galloyl	Galloylated flavan-3-ols	*Vitis vinifera*		Terrier *et al.* (2009)
VvGAT2					
ECGT	Galloyl	Galloylated flavan-3-ols	*Camellia sinensis*		Liu *et al.* (2012)
SCPL17	Benzoyl or sinapoyl	Benzoyl or sinapoyl glucosinolates	*Arabidopsis thaliana*		Lee *et al.* (2012)
AA7G-AT	Hydoxybenzoyl	Anthocyanins	*Delphinium grandiflorum*	Vacuole	Nishizaki *et al.* (2013)
AA7GBG-AT					

In bold letters: characterized SCPL-ATs; *, enzymes that do not use phenolic compounds; –, no name.
Abbreviations and GenBank accession numbers are: GAC, glucose acyltransferase (AAF64227); CtAT1, *Clitoria ternalea* acyltransferase 1 (BAF99695, Patent Publication no. WO/2007/046148); SMT, 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase (AAF78760); SCT, 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase (AAK52316); BnSCT1, *Brassica napus* 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase 1 (AAQ91191); BnSCT2 (CAM91991); SAT, 1-*O*-sinapoyl-β-glucose:anthocyanin sinapoyltransferase (AEC07395); SST, 1-*O*-sinapoyl-β-glucose:1-*O*-sinapoyl-β-glucose sinapoyltransferase (AEC07397); SCPL1 = SAD7, saponin-deficient 7 (ACT21078); AMalT, anthocyanin malyltransferase; OsBISCPL1, *Oryza sativa* BTH-induced serine carboxypeptidase 1; DkSCPL1, *Diospyros kaki* Serine Carboxypeptidase Like 1 (BAF56655); DkSCPL2 (BAH89272); VvGAT1, *Vitis vinifera* glucose-acyltransferase 1; VvGAT2, *Vitis vinifera* glucose-acyltransferase 2; ECGT, epicatechin:1-*O*-galloyl-β-D-glucose *O*-galloyltransferase; SCPL17 (AAS99709); AA7G-AT, acyl-Glc-dependent anthocyanin 7-*O*-glucoside acyltransferase; AA7GBG-AT, acyl-Glc-dependent anthocyanin 7-*O*-(6-*O*-(4-*O*-(glucosyl)-oxybenzoyl)-glucoside) acyltransferase.

D-glucose is also used as a donor molecule by sinapoylglucose: sinapoylglucose sinapoyltransferase (SST) to synthesize 1,2-di-*O*-sinapoyl-β-D-glucose, and by sinapoylglucose:anthocyanin sinapoyltransferase (SAT) to synthesize sinapoylated anthocyanins (Fraser *et al.*, 2007). In *Clitoria ternalea*, CtAT1 coumaroylates an anthocyanin called ternatin C5 (delphinidin 3-*O*-(6″-*O*-malonyl)-β-glucoside-3′,5′-di-*O*-β-glucoside) to produce ternatin C3 (Noda *et al.*, 2007, Patent WO/2007/046148). Mugford *et al.* (2009) identified SCPL1, an AT that is able to synthesize a triterpene glycoside (avenacin) in *Avena strigosa*. SCPL1 can acylate des-acyl avenacin using *N*-methyl anthraniloyl-*O*-glucose as the substrate (Fig. 4).

1-*O*-β-glucose ester dependent and not yet characterized ATs

Only nine SCPL-AT genes have been fully characterized. However, several studies have reported 1-*O*-β-glucose ester-dependent acylation using classical enzymatic approaches or have identified putative SCPL-AT genes through transcriptomic or genetic screening (Table 2).

An enzyme purified from *Zea mays* is responsible for indole acetyl-*myo*-inositol biosynthesis (Kowalczyk *et al.*, 2003). This enzyme uses an indoleacetyl moiety from 1-*O*-(indole-3-acetyl)-β-D-glucose to acylate *myo*-inositol. A recent study has demonstrated the transfer of an indoleacetyl moiety to different saccharides catalyzed by an enzyme purified from *Z. mays* endosperm (Starzyńska & Kowalczyk, 2012). An inverse genetic approach in Arabidopsis suggests that SCPL17 could acylate hydroxylated glucosinolates with 1-*O*-sinapoyl-β-D-glucose to form sinapoylated glucosinolates (Lee *et al.*, 2012). The acylation of cyanidin triglucoside (cyanidin 3-*O*-(2″-*O*-xylosyl)-6″-*O*-glucosylgalactoside) by hydroxycinnamoyl moieties from the 1-*O*-β-glucose ester of HCAs has been observed with enzyme extracts of *Daucus carota* cells (Glaessgen & Seitz, 1992). Anthocyanin malyl transferase (AMalT) activity has been identified in crude enzyme extracts of *Dianthus caryophyllus* using the 1-*O*-β-glucose ester of malate as the donor molecule (Abe *et al.*, 2008). In *Delphinium grandiflorum* crude enzyme extracts, the enzymatic activity of two anthocyanin *p*-benzoyltransferases (AA7G-AT and AA7GBG-AT) that use *p*-hydroxybenzoyl-glucose as the acyl donor have been detected (Nishizaki *et al.*, 2013). These two enzymes take part in the step-by-step biosynthesis of viodelphin.

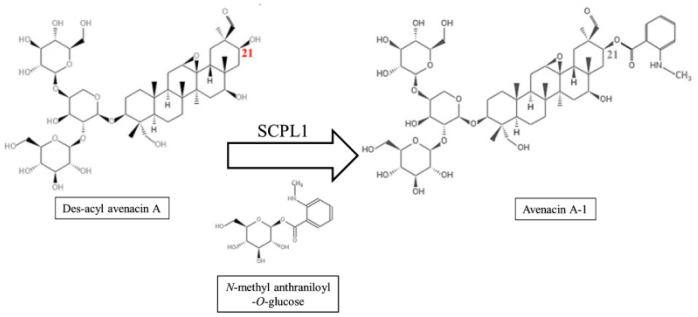


Fig. 4 Acylation reaction catalyzed by the SCPL acyltransferase SCPL1 from the monocot *Avena strigosa* (Mugford *et al.*, 2009). The *N*-methyl anthraniloyl moiety from *N*-methyl anthraniloyl-*O*-glucose is esterified with the hydroxyl of the carbon C21 of des-acyl avenacin A to produce avenacin A-1.

Gallotannin biosynthesis is thought to be linked to transacylation from β-glucogallin, a 1-*O*-β-glucose ester of gallic acid. Indeed, several enzymatic activities have been observed in *Quercus* species (for a review, see Niemetz & Gross, 2005). To date, no gene coding for ATs that are able to catalyze such reactions has been characterized. β-Glucogallin could be a precursor in the biosynthesis of another type of tannin, the proanthocyanidins, which are notably responsible for fruit, tea and wine astringency. In certain species, such as *Camellia sinensis*, *V. vinifera* and *Diospyros kaki*, flavan-3-ols, including proanthocyanidin subunits, can be acylated with a galloyl moiety. Flavan-3-ol galloylation activity has been observed recently in crude extracts of tea leaves that are known to be rich in galloylated flavan-3-ols (Liu *et al.*, 2012). However, the gene responsible for this reaction has not been isolated.

Several studies have shown that transcription factors can orchestrate the expression of putative 'decoration' genes, in addition to genes involved in the biosynthesis of a compound *sensu stricto*. In *A. thaliana*, *SCPL17* has been identified by Lee *et al.* (2012) from microarray data as a gene regulated by a MYB transcription factor, which controls glucosinolate biosynthesis. Studies identifying genes in *D. kaki* that are differentially expressed in astringent and non-astringent fruit revealed two SCPL genes (*DkSCPL1* and *DkSCPL2*) (Ikegami *et al.*, 2007; Akagi *et al.*, 2009). Several SCPL genes are up-regulated by MYB transcription factors controlling proanthocyanidin biosynthesis in grapes (*VvGAT1* and *VvGAT2*; Terrier *et al.*, 2009; Carrier *et al.*, 2013; Koyama *et al.*, 2014). These enzymes in persimmons and grapes may catalyze the proanthocyanidin galloylation described above.

Most SCPL-ATs use 1-*O*-β-glucose esters of phenolic acids as substrates, but 1-*O*-β-glucose esters of aliphatic acids may also be used as substrates. For example, GAC in wild tomato uses *O*-isobutyrylglucose, and AMalT in *Dianthus caryophyllus* uses the 1-*O*-β-glucose ester of malate.

Two-step SCPL acylation: a localization puzzle

The presence of an N-terminal signal peptide on immunolocalization can suggest vacuolar localization of SCPL-ATs. Indeed, SCPL-AT proteins have been detected in the vacuole (Hause *et al.*, 2002; Noda *et al.*, 2007; Mugford *et al.*, 2013; Fig. 3). The formation of 1-*O*-β-glucose esters is catalyzed by specific UDP-glucosyltransferases (UGTs) (Li *et al.*, 2001). Notably, UDP-Glc:sinapic acid glucosyltransferase (SGT) is able to synthesize sinapoylglucose, which is used by several SCPL-ATs in Brassicaceae (Milkowski *et al.*, 2000). In *V. vinifera*, three UGTs have recently been shown to form hydroxycinnamoyl glucose esters and β-glucogallin (Khater *et al.*, 2012). These UGTs are thought to be localized in the cytosolic compartment (Achnine *et al.*, 2005; Mugford *et al.*, 2013; Fig. 3).

Thus, UGT and SCPL-AT activity can occur in two different compartments. As a result, vacuolar activity of the SCPL protein suggests that both 1-*O*-β-glucose esters and acyl acceptors are imported into the vacuole. Indeed, glycosylated compounds are found in the vacuolar compartment in which acylation may occur (Vogt & Jones, 2000). Malate, which is a potential acyl acceptor, can be imported directly into the vacuole via a tonoplastic channel in *A. thaliana* (AtALMT9, Kovermann *et al.*, 2007; Fig. 3).

Glycosylation is often an important step in molecule transport. In *M. truncatula*, UGT72L1 synthesizes epicatechin 3′-*O*-glucoside (Pang *et al.*, 2008), a form of flavan-3-ol that is imported into the vacuole (Zhao & Dixon, 2009b). Recently, in *V. vinifera*, ABCC1 was identified as an ATP Binding Cassette (ABC) protein that allows for the import of glycosylated anthocyanins into the vacuole (Francisco *et al.*, 2013). In addition, glucose-ester can also be transported through endomembranes. Indeed, both ABC-type and proton gradient-driven transporters have been implicated recently in the vacuolar uptake of abscisic acid glucose-ester (Burla *et al.*, 2013). Hence, glycosylation fulfills a double function. In addition to providing an activated substrate for SCPL-ATs, glycosylation adds a glucose tag, which is often compulsory for recognition by transporters.

Conclusion

The diversity and complexity of secondary metabolites in vascular plants contribute to the capacity of these sessile organisms for adaptation. Acylation contributes to such diversity. Phenolic compounds are frequently used as acyl donors and/or acceptors for acylations catalyzed by BAHD- and SCPL-ATs.

The identification of numerous genes in the *A. thaliana* genome coding for both types of ATs indicates no species specificity for these ATs. Both AT types use HCA esters to acylate a wide range of molecules (e.g. flavonoids, phytoalexins, quinic and shikimic acid, choline), and use non-phenolic compounds as donor (e.g.

malonate, isobutyrate) or acceptor (e.g. glucosinolates, malate) molecules. To date, acylation with HBAs has only been reported with hydroxybenzoyl-glucose, but not hydroxybenzoyl-CoA, suggesting the involvement of putative SCPL-ATs rather than BAHD-ATs. Both AT types can produce compounds belonging to the same family, in the same species (e.g. *A. thaliana* anthocyanin acylation by aromatic BAHD- and SCPL-ATs) or in different species (CGA biosynthesis in *Ipomoea batatas* and *Nicotiana tabacum*).

The main difference resides in their localization: BAHD-ATs are cytosolic, whereas SCPL-ATs are vacuolar. BAHD-ATs target anthocyanins for vacuolar storage, whereas SCPL-ATs regulate downstream modifications in the vacuole. SCPL-AT activity in the vacuole requires the presence of different upstream actors. Indeed, 1-*O*-β-glucose esters are first produced by cytosolic UGTs, and then routed to the vacuole via specific tonoplastic transporters. The existence and localization of upstream actors (enzymes producing acylating and acylated molecules, and their relative transporters) could govern the recruitment of one of the AT types. The molecules imported into the vacuole following BAHD-AT acylation could be re-acylated by SCPL-AT. Hence, both AT types may be recruited in different compartments for concomitant activities, synergistically promoting the accumulation of metabolites. In addition, the biosynthesis of acylated defense compounds in the cytosol is potentially toxic for the plant, and thus vacuolar SCPL-AT activity may offer an alternative pathway to cope with this scenario. Such an alternative pathway would allow for rapid and efficient synthesis and vacuolar trapping of acylated compounds and avoid cytosolic synthesis before tonoplastic transport.

In any case, the determination of the factors that govern the involvement of either AT type for acylation remains a challenge for future research. The comparison of BAHD- and SCPL-AT expression patterns, with metabolite profiles in different developmental stages, tissues and environmental conditions, could help decipher the respective roles of these ATs. The characterization of new members of the SCPL-AT family opens up the way for new lines of research.

References

Abe Y, Tera M, Sasaki N, Okamura M, Umemoto N, Momose M, Kawahara N, Kamakura H, Goda Y, Nagasawa K *et al.* 2008. Detection of 1-*O*-malylglucose: pelargonidin 3-*O*-glucose-6″-*O*-malyltransferase activity in carnation (*Dianthus caryophyllus*). *Biochemical and Biophysical Research Communications* **373**: 473–477.

Achnine L, Huhman DV, Farag MA, Sumner LW, Blount JW, Dixon RA. 2005. Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. *Plant Journal* **41**: 875–887.

Agrimi G, Russo A, Pierri CL, Palmieri F. 2012. The peroxisomal NAD$^+$ carrier of *Arabidopsis thaliana* transports coenzyme A and its derivatives. *Journal of Bioenergetics and Biomembranes* **44**: 333–340.

Akagi T, Ikegami A, Suzuki Y, Yoshida J, Yamada M, Sato A, Yonemori K. 2009. Expression balances of structural genes in shikimate and flavonoid biosynthesis cause a difference in proanthocyanidin accumulation in persimmon (*Diospyros kaki* Thunb.) fruit. *Planta* **230**: 899–915.

Bartley LE, Peck ML, Kim S-R, Ebert B, Manisseri C, Chiniquy DM, Sykes R, Gao L, Rautengarten C, Vega-Sánchez ME *et al.* 2013. Overexpression of a BAHD acyltransferase, OsAt10, alters rice cell wall hydroxycinnamic acid content and saccharification. *Plant Physiology* **161**: 1615–1633.

Bassard JE, Ullmann P, Bernier F, Werck-Reichhart D. 2010. Phenolamides: bridging polyamines to the phenolic metabolism. *Phytochemistry* **71**: 1808–1824.

Berger A, Meinhard J, Petersen M. 2006. Rosmarinic acid synthase is a new member of the superfamily of BAHD acyltransferases. *Planta* **224**: 1503–1510.

Boatright J, Negre F, Chen X, Kish CM, Wood B, Peel G, Orlova I, Gang D, Rhodes D, Dudareva N. 2004. Understanding *in vivo* benzenoid metabolism in petunia petal tissue. *Plant Physiology* **135**: 1993–2011.

Burhenne K, Kristensen BK, Rasmussen SK. 2003. A new class of *N*-hydroxycinnamoyltransferases. Purification, cloning, and expression of a barley agmatine coumaroyltransferase (EC 23. 1. 64). *Journal of Biological Chemistry* **278**: 13919–13927.

Burla B, Pfrunder S, Nagy R, Francisco RM, Lee Y, Martinoia E. 2013. Vacuolar transport of abscisic acid glucosyl ester is mediated by ATP-binding cassette and proton-antiport mechanisms in Arabidopsis. *Plant Physiology* **163**: 1446–1458.

Carrier G, Huang YF, Le Cunff L, Fournier-Level A, Vialet S, Souquet JM, Cheynier V, Terrier N, This P. 2013. Selection of candidate genes for grape proanthocyanidin pathway by an integrative approach. *Plant Physiology and Biochemistry* **72**: 87–95.

Chedgy RJ, Köllner TG, Constabel CP. 2015. Functional characterization of two acyltransferases from *Populus trichocarpa* capable of synthesizing benzyl benzoate and salicyl benzoate, potential intermediates in salicinoid phenolic glycoside biosynthesis. *Phytochemistry* **113**: 149–159.

Cheng AX, Gou JY, Yu XH, Yang H, Fang X, Chen XY, Liu CJ. 2013. Characterization and ectopic expression of a *Populus* hydroxyacid hydroxycinnamoyltransferase. *Molecular Plant* **6**: 1889–1903.

Cheynier V, Comte G, Davies KM, Lattanzio V, Martens S. 2013. Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology. *Plant Physiology and Biochemistry* **72**: 1–20.

Clé C, Hill LM, Niggeweg R, Martin CR, Guisez Y, Prinsen E, Jansen MA. 2008. Modulation of chlorogenic acid biosynthesis in *Solanum lycopersicum*; consequences for phenolic accumulation and UV-tolerance. *Phytochemistry* **69**: 2149–2156.

Comino C, Lanteri S, Portis E, Acquadro A, Romani A, Hehn A, Larbat R, Bourgaud F. 2007. Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. *BMC Plant Biology* **20**: 7–14.

Dahlbender B, Strack D. 1984. Nitrogen nutrition and the accumulation of free and sinapoyl-bound malic acid in *Raphanus sativus* cotyledons. *Planta* **161**: 142–147.

D'Auria JC. 2006. Acyltransferases in plants: a good time to be BAHD. *Current Opinion in Plant Biology* **9**: 331–340.

D'Auria JC, Chen F, Pichersky E. 2002. Characterization of an acyltransferase capable of synthesizing benzylbenzoate and other volatile esters in flowers and damaged leaves of *Clarkia breweri*. *Plant Physiology* **130**: 466–476.

D'Auria JC, Reichelt M, Luck K, Svatos A, Gershenzon J. 2007. Identification and characterization of the BAHD acyltransferase malonyl CoA:anthocyanidin 5-*O*-glucoside-6″-*O*-malonyltransferase (At5MAT) in *Arabidopsis thaliana*. *FEBS Letters* **581**: 872–878.

Dhaubhadel S, Farhangkhoee M, Chapman R. 2008. Identification and characterization of isoflavonoid specific glycosyltransferase and malonyltransferase from soybean seeds. *Journal of Experimental Botany* **59**: 981–994.

Dudareva N, D'Auria JC, Nam KH, Raguso RA, Pichersky E. 1998. Acetyl-CoA: benzylalcohol acetyltransferase – an enzyme involved in floral scent production in *Clarkia breweri*. *Plant Journal* **14**: 297–304.

Escamilla-Treviño LL, Shen H, Hernandez T, Yin Y, Xu Y, Dixon RA. 2014. Early lignin pathway enzymes and routes to chlorogenic acid in switchgrass (*Panicum virgatum* L.). *Plant Molecular Biology* **84**: 565–576.

Facchini PJ, Hagel J, Zulak KG. 2002. Hydroxycinnamic acid amide metabolism: physiology and biochemistry. *Canadian Journal of Botany* **80**: 577–589.

Feucht W, Dithmar H, Polster J. 2004. Nuclei of tea flowers as targets for flavanols. *Plant Biology* **6**: 696–701.

Francisco RM, Regalado A, Ageorges A, Burla BJ, Bassin B, Eisenach C, Zarrouk O, Vialet S, Marlin T, Chaves MM *et al.* 2013. ABCC1, an ATP binding cassette

protein from grape berry, transports anthocyanidin 3-*O*-glucosides. *Plant Cell* **25**: 1840–1854.

Fraser CM, Rider LW, Chapple C. 2005. An expression and bioinformatics analysis of the Arabidopsis serine carboxypeptidaselike gene family. *Plant Physiology* **138**: 1136–1148.

Fraser CM, Thompson MG, Shirley AM, Ralph J, Schoenherr JA, Sinlapadech T, Hall MC, Chapple C. 2007. Related Arabidopsis serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. *Plant Physiology* **144**: 1986–1999.

Fujiwara H, Tanaka Y, Yonekura-Sakakibara K, Fukuchi-Mizutani M, Nakao M, Fukui Y, Yamaguchi M, Ashikari T, Kusumi T. 1998. cDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from *Gentiana triflora*. *Plant Journal* **16**: 421–431.

Glaessgen WE, Seitz HU. 1992. Acylation of anthocyanins with hydroxycinnamic acids via 1-O-acylglucosides by protein preparations from cell cultures of *Daucus carota* L. *Planta* **186**: 582–585.

Gomez C, Conejero G, Torregrosa L, Cheynier V, Terrier N, Ageorges A. 2011. *In vivo* grapevine anthocyanin transport involves vesicle mediated trafficking and the contribution of anthoMATE transporters and GST. *Plant Journal* **67**: 960–970.

Gomez C, Terrier N, Torregrosa L, Vialet S, Fournier-Level A, Verriès C, Souquet JM, Mazauric JP, Klein M, Cheynier V *et al.* 2009. Grapevine MATE-type proteins act as vacuolar H^+-dependent acylated anthocyanin transporters. *Plant Physiology* **150**: 402–415.

Gou JY, Yu XH, Liu CJ. 2009. A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in Arabidopsis. *Proceeding of the National Academy of Sciences, USA* **106**: 18855–18860.

Grienenberger E, Besseau S, Geoffroy P, Debayle D, Heintz D, Lapierre C, Pollet B, Heitz T, Legrand M. 2009. A BAHD acyltransferase is expressed in the tapetum of Arabidopsis anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. *Plant Journal* **58**: 246–259.

Gross GG. 1983. Synthesis of mono-, di- and trigalloyl-β-D-glucose by β-glucogallin-dependent galloyltransferases from oak leaves. *Zeitschrift für Naturforschung* **38c**: 519–523.

Harbone JB, Williams CA. 2000. Advances in flavonoid research since 1992. *Phytochemistry* **55**: 481–504.

Hause B, Meyer K, Viitanen PV, Chapple C, Strack D. 2002. Immunolocalization of 1-*O*-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. *Planta* **215**: 26–32.

Heller W, Forkmann G. 1994. Biosynthesis of flavonoids. In: Harborne JB, ed. *The flavonoids, advances in research since 1986*. London, UK: Chapman & Hall, 499–535.

Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, Legrand M. 2004. Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *Plant Cell* **16**: 1446–1465.

Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, Legrand M. 2005. Acyltransferase-catalysed *p*-coumarate ester formation is a committed step of lignin biosynthesis. *Plant Biosystems* **139**: 50–53.

Hoffmann L, Maury S, Martz F, Geoffroy P, Legrand M. 2003. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *Journal of Biological Chemistry* **278**: 95–103.

Ikegami A, Eguchi S, Kitajima A, Inoue K, Yonemori K. 2007. Identification of genes involved in proanthocyanidin biosynthesis of persimmon (*Diospyros kaki*) fruit. *Plant Science* **172**: 1037–1047.

Khater F, Fournand D, Vialet S, Meudec E, Cheynier V, Terrier N. 2012. Identification and functional characterization of cDNAs coding for hydroxybenzoate/hydroxycinnamate glucosyltransferases co-expressed with genes related to proanthocyanidin biosynthesis. *Journal of Experimental Botany* **63**: 1201–1214.

Kim BG, Kim IA, Ahn JH. 2011. Characterization of hydroxycinnamoyl-coenzyme A shikimate hydroxycinnamoyltransferase from *Populus euramericana*. *Journal of the Korean Society for Applied Biological Chemistry* **54**: 817–821.

Kim IA, Kim BG, Kim M, Ahn JH. 2012. Characterization of hydroxycinnamoyltransferase from rice and its application for biological synthesis of hydroxycinnamoyl glycerols. *Phytochemistry* **76**: 25–31.

Kojima M, Uritani I. 1972. Elucidation of the structure of a possible intermediate in chlorogenic acid biosynthesis in sweet potato root tissue. *Plant and Cell Physiology* **13**: 1075–1084.

Kosma DK, Molina I, Ohlrogge JB, Pollard M. 2012. Identification of an Arabidopsis fatty alcohol:caffeoyl-coenzyme A acyltransferase required for the synthesis of alkyl hydroxycinnamates in root waxes. *Plant Physiology* **160**: 237–248.

Kovermann P, Meyer S, Hörtensteiner S, Picco C, Scholz-Starke J, Ravera S, Lee Y, Martinoia E. 2007. The Arabidopsis vacuolar malate channel is a member of the ALMT family. *Plant Journal* **52**: 1169–1180.

Kowalczyk S, Jakubowska A, Zielinska E, Bandurski RS. 2003. Bifunctional indole-3-acetyl transferase catalyses synthesis and hydrolysis of indole-3-acetyl-myo-inositol in immature endosperm of *Zea mays*. *Plant Physiology* **119**: 165–174.

Koyama K, Numata M, Nakajima I, Goto-Yamamoto N, Matsumura H, Tanaka N. 2014. Functional characterization of a new grapevine MYB transcription factor and regulation of proanthocyanidin biosynthesis in grapes. *Journal of Experimental Botany* **65**: 4433–4449.

Lallemand LA, Zubieta C, Lee SG, Wang Y, Acajjaoui S, Timmins J, McSweeney S, Jez JM, McCarthy JG, McCarthy AA. 2012. A structural basis for the biosynthesis of the major chlorogenic acids found in coffee. *Plant Physiology* **160**: 249–260.

Landmann C, Hücherig S, Fink B, Hoffmann T, Dittlein D, Coiner HA, Schwab W. 2011. Substrate promiscuity of a rosmarinic acid synthase from lavender (*Lavandula angustifolia* L.). *Planta* **234**: 305–320.

Landry LG, Chapple CCS, Last RL. 1995. Arabidopsis mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. *Plant Physiology* **109**: 1159–1166.

Lee S, Kaminaga Y, Cooper B, Pichersky E, Dudareva N, Chapple C. 2012. Benzoylation and sinapoylation of glucosinolate R-groups in Arabidopsis. *Plant Journal* **72**: 411–422.

Lehfeldt C, Shirley AM, Meyer K, Ruegger MO, Cusumano JC, Viitanen PV, Strack D, Chapple C. 2000. Cloning of the SNG1 gene of Arabidopsis reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. *Plant Cell* **12**: 1295–1306.

Lepelley M, Cheminade G, Tremillon N, Simkin A, Caillet V, McCarthy J. 2007. Chlorogenic acid synthesis in coffee: an analysis of CGA content and real-time RT-PCR expression of HCT, HQT, C3H1, and CCoAOMT1 genes during grain development in *C. canephora*. *Plant Science* **172**: 978–996.

Li Y, Baldauf S, Lim EK, Bowles DJ. 2001. Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. *Journal of Biological Chemistry* **276**: 4338–4343.

Li AX, Steffens JC. 2000. An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. *Proceedings of the National Academy of Sciences, USA* **97**: 6902–6907.

Liu Y, Gao L, Liu L, Yang Q, Lu Z, Nie Z, Wang Y, Xia T. 2012. Purification and characterization of a novel galloyltransferase involved in catechin galloylation in the tea plant (*Camellia sinensis*). *Journal of Biological Chemistry* **287**: 44406–44417.

Luo J, Fuell C, Parr A, Hill L, Bailey P, Elliott K, Fairhurst SA, Martin C, Michael AJ. 2009. A novel polyamine acyltransferase responsible for the accumulation of spermidine conjugates in Arabidopsis seed. *Plant Cell* **21**: 318–333.

Luo J, Nishiyama Y, Fuell C, Taguchi G, Elliott K, Hill L, Tanaka Y, Kitayama M, Yamazaki M, Bailey P *et al.* 2007. Convergent evolution in the BAHD family of acyl transferases: identification and characterization of anthocyanin acyl transferases from *Arabidopsis thaliana*. *Plant Journal* **50**: 678–695.

Milkowski C, Baumert A, Schmidt D, Nehlin L, Strack D. 2004. Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. *Plant Journal* **38**: 80–92.

Milkowski C, Baumert A, Strack D. 2000. Cloning and heterologous expression of a rape cDNA encoding UDP-glucose:sinapate glucosyltransferase. *Planta* **211**: 883–886.

Milkowski C, Strack D. 2004. Serine carboxypeptidase-like acyltransferases. *Phytochemistry* **65**: 517–524.

Moglia A, Lanteri S, Comino C, Hill L, Knevitt D, Cagliero C, Rubiolo P, Bornemann S, Martin C. 2014. Dual catalytic activity of hydroxycinnamoyl-Coenzyme A quinate transferase from tomato allows it to moonlight in the synthesis of both mono- and dicaffeoylquinic acids. *Plant Physiology* **166**: 1777–1787.

Molina I, Li-Beisson Y, Beisson F, Ohlrogge JB, Pollard M. 2009. Identification of an Arabidopsis feruloyl-coenzyme A transferase required for suberin synthesis. *Plant Physiology* **151**: 1317–1328.

Mugford ST, Louveau T, Melton R, Qi X, Bakht S, Hill L, Tsurushima T, Honkanen S, Rosser SJ, Lomonossoff GP *et al.* 2013. Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat. *Plant Cell* **25**: 1078–1092.

Mugford ST, Milkowski C. 2012. Serine carboxypeptidase-like acyltransferases from plants. *Methods in Enzymology* **516**: 279–297.

Mugford ST, Qi X, Bakht S, Hill L, Wegel E, Hughes RK, Papadopoulou K, Melton R, Philo M, Sainsbury F *et al.* 2009. A Serine Carboxypeptidase-Like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. *Plant Cell* **21**: 2473–2484.

Muroi A, Ishihara A, Tanaka C, Ishizuka A, Takabayashi J, Miyoshi H, Nishioka T. 2009. Accumulation of hydroxycinnamoyl acid amides induced by pathogen infection and identification of agmatine coumaroyltransferase in *Arabidopsis thaliana*. *Planta* **230**: 517–527.

Nakayama T, Suzuki H, Nishino T. 2003. Anthocyanin acyltransferases: specificities, mechanism, phylogenetics, and applications. *Journal of Molecular Catalysis B: Enzymatic* **23**: 117–132.

Niemetz R, Gross GG. 2005. Enzymology of gallotannin and ellagitannin biosynthesis. *Phytochemistry* **66**: 2001–2011.

Niggeweg R, Michael AJ, Martin C. 2004. Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nature Biotechnology* **22**: 746–754.

Nishizaki Y, Yasunaga M, Okamoto E, Okamoto M, Hirose Y, Yamaguchi M, Ozeki Y, Sasaki N. 2013. *p*-Hydroxybenzoyl-glucose is a zwitter donor for the biosynthesis of 7-polyacylated anthocyanin in Delphinium. *Plant Cell* **25**: 4150–4165.

Noda N, Higaki T, Sano T, Kazuma K, Sasaki T, Hasegawa S, Suzuki M. 2007. The structural localization and substrate specificity of 1-O-acylglucose dependent anthocyanin acyltransferase from blue petals of butterfly pea (*Clitoria ternatea*). Abstracts of the 48th Annual Meeting of the Japanese Society of Plant Physiologists, March 28–30, 2007, Matsuyama, Japan. *Plant and Cell Physiology.* **48**: S222.

Onkokesung N, Gaquerel E, Kotkar H, Kaur H, Baldwin IT, Galis I. 2012. MYB8 controls inducible phenolamide levels by activating three novel hydroxycinnamoyl-coenzyme A:polyamine transferases in *Nicotiana attenuata*. *Plant Physiology* **158**: 389–407.

Pang Y, Peel GJ, Sharma SB, Tang Y, Dixon RA. 2008. A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of *Medicago truncatula*. *Proceedings of the National Academy of Sciences, USA* **105**: 14210–14215.

Panikashvili D, Shi JX, Schreiber L, Aharoni A. 2009. The Arabidopsis DCR encoding a soluble BAHD acyltransferase is required for cutin polyester formation and seed hydration properties. *Plant Physiology* **151**: 1773–1789.

Peng X, Li W, Wang W, Bai G. 2010. Cloning and characterization of a cDNA coding a hydroxycinnamoyl-CoA quinate hydroxycinnamoyl transferase involved in chlorogenic acid biosynthesis in *Lonicera japonica*. *Planta Medica* **76**: 1921–1926.

Petersen M, Abdullah Y, Benner J, Eberle D, Gehlen K, Hücherig S, Janiak V, Kim KH, Sander M, Weitzel C *et al.* 2009. Evolution of rosmarinic acid biosynthesis. *Phytochemistry* **70**: 1663–1679.

Petrik DL, Karlen SD, Cass CL, Padmakshan D, Lu F, Liu S, Le Bris P, Antelme S, Santoro N, Wilkerson CG *et al.* 2014. *p*-Coumaroyl-CoA:monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in *Brachypodium distachyon*. *Plant Journal* **77**: 713–726.

Pollard M, Beisson F, Li Y, Ohlrogge JB. 2008. Building lipid barriers: biosynthesis of cutin and suberin. *Trends in Plant Science* **13**: 236–246.

Rautengarten C, Ebert B, Ouellet M, Nafisi M, Baidoo EE, Benke P, Stranne M, Mukhopadhyay A, Keasling JD, Sakuragi Y *et al.* 2012. Arabidopsis deficient in cutin ferulate encodes a transferase required for feruloylation of ω-hydroxy fatty acids in cutin polyester. *Plant Physiology* **158**: 654–665.

Sander M, Petersen M. 2011. Distinct substrate specificities and unusual substrate flexibilities of two hydroxycinnamoyltransferases, rosmarinic acid synthase and hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl-transferase, from *Coleus blumei* Benth. *Planta* **233**: 1157–1171.

Schmidt GW, Jirschitzka J, Porta T, Reichelt M, Luck K, Torre JC, Dolke F, Varesio E, Hopfgartner G, Gershenzon J *et al.* 2015. The last step in cocaine biosynthesis is catalyzed by a BAHD acyltransferase. *Plant Physiology* **167**: 89–101.

Serra O, Hohn C, Franke R, Prat S, Molinas M, Figueras M. 2010. A feruloyl transferase involved in the biosynthesis of suberin and suberin-associated wax is required for maturation and sealing properties of potato periderm. *Plant Journal* **62**: 277–290.

Shirley AM, Chapple C. 2003. Biochemical characterization of sinapoyl:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. *Journal of Biological Chemistry* **278**: 19870–19877.

Shockey J, Browse J. 2011. Genome-level and biochemical diversity of the acyl-activating enzyme superfamily in plants. *Plant Journal* **66**: 143–160.

Sonnante G, D'Amore R, Blanco E, Pierri CL, De Palma M, Luo J, Tucci M, Martin C. 2010. Novel hydroxycinnamoyl-coenzyme A quinate transferase genes from artichoke are involved in the synthesis of chlorogenic acid. *Plant Physiology* **153**: 1224–1238.

Starzyńska E, Kowalczyk S. 2012. Novel 1-*O*-indole-3-acetyl-β-D-glucose-dependent acyltransferase transferring indoleacetyl moiety to some mono- di- and oligosaccharides. *Acta Physiologiae Plantarum* **34**: 53–63.

Stehle F, Stubbs MT, Strack D, Milkowski C. 2008. Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism. *FEBS Journal* **275**: 775–787.

St-Pierre B, De Luca V. 2000. Evolution of acyltransferase genes: origin and diversification of the BAHD superfamily of acyltransferases involved in secondary metabolism. In: John RI, Romeo T, Varin L, De Luca V, eds. *Recent advances in phytochemistry evolution of metabolic pathways*. Oxford, UK: Elsevier Science, 285–315.

St-Pierre B, Laflamme P, Alarco AM, De Luca V. 1998. The terminal *O*-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A dependent acyl transfer. *Plant Journal* **14**: 703–713.

Sullivan ML. 2009. A novel red clover hydroxycinnamoyl transferase has enzymatic activities consistent with a role in phaselic acid biosynthesis. *Plant Physiology* **150**: 1866–1879.

Suzuki H, Nakayama T, Nagae S, Yamaguchi M, Iwashita T, Fukui Y, Nishino T. 2004a. cDNA cloning and functional characterization of flavonol 3-*O*-glucoside-6″-*O*-malonyltransferases from flowers of *Verbena hybrida* and *Lamium purpureum*. *Journal of Molecular Catalysis B: Enzymatic* **28**: 87–93.

Suzuki H, Nakayama T, Yamaguchi MA, Nishino T. 2004b. cDNA cloning and characterization of two *Dendranthema* × *morifolium* anthocyanin malonyltransferases with different functional activities. *Plant Science* **166**: 89–96.

Suzuki H, Nakayama T, Yonekura-Sakakibara K, Fukui Y, Nakamura N, Nakao M, Tanaka Y, Yamaguchi M, Kusumi T, Nishino T. 2001. Malonyl-CoA: anthocyanin 5-*O*-glucoside-6″-*O*-malonyltransferase from scarlet sage (*Salvia splendens*) flowers. *Journal of Biological Chemistry* **276**: 49013–49019.

Suzuki H, Nakayama T, Yonekura-Sakakibara K, Fukui Y, Nakamura N, Yamaguchi M, Tanaka Y, Kusumi T, Nishino T. 2002. cDNA cloning, heterologous expressions, and functional characterization of malonyl-coenzyme A:anthocyanidin 3-*O*-glucoside-6″-*O*-malonyltransferase from *Dahlia* flowers. *Plant Physiology* **130**: 2142–2151.

Suzuki H, Nishino T, Nakayama T. 2007. cDNA cloning of a BAHD acyltransferase from soybean (*Glycine max*): isoflavone 7-*O*-glucoside-6″-*O*-malonyltransferase. *Phytochemistry* **68**: 2035–2042.

Suzuki H, Sawada S, Watanabe K, Nagae S, Yamaguchi M, Nakayama T, Nishino T. 2004c. Identification and characterization of a novel anthocyanin malonyltransferase from scarlet sage (*Salvia splendens*) flowers: an enzyme that is phylogenetically separated from other anthocyanin acyltransferases. *Plant Journal* **38**: 994–1003.

Suzuki H, Sawada S, Yonekura-Sakakibara K, Nakayama T, Yamaguchi M, Nishino T. 2003. Identification of a cDNA encoding malonyl-CoenzymeA:

anthocyanidin 3-*O*-glucoside 6″-*O*-malonyltransferase from cineraria (*Senecia curentus*) flowers. *Plant Biotechnology* **20**: 229–234.

Taguchi G, Shitchi Y, Shirasawa S, Yamamoto H, Hayashida N. 2005. Molecular cloning, characterization, and downregulation of an acyltransferase that catalyzes the malonylation of flavonoid and naphthol glucosides in tobacco cells. *Plant Journal* **42**: 481–491.

Taguchi G, Ubukata T, Nozue H, Kobayashi Y, Takahi M, Yamamoto H, Hayashida N. 2010. Malonylation is a key reaction in the metabolism of xenobiotic phenolic glucosides in Arabidopsis and tobacco. *Plant Journal* **63**: 1031–1041.

Terrier N, Torregrosa L, Ageorges A, Vialet S, Verries C, Cheynier V, Romieu C. 2009. Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. *Plant Physiology* **149**: 1028–1041.

Tuominen LK, Johnson VE, Tsai CJ. 2011. Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among Populus paralogues. *BMC Genomics* **12**: 236.

Unno H, Ichimaida F, Suzuki H, Takahashi S, Tanaka Y, Saito A, Nishino T, Kusunoki M, Nakayama T. 2007. Structural and mutational studies of anthocyanin malonyltransferases establish the features of BAHD enzyme catalysis. *Journal of Biological Chemistry* **282**: 15812–15822.

Varbanova M, Porter K, Lu F, Ralph J, Hammerschmidt R, Jones AD, Day B. 2011. Molecular and biochemical basis for stress-induced accumulation of free and bound *p*-coumaraldehyde in cucumber. *Plant Physiology* **157**: 1056–1066.

Villegas RJA, Kojima M. 1986. Purification and characterization of hydroxycinnamoyl D-glucose quinate hydroxycinnamoyltransferase in the root of sweet potato, *Ipomoea batatas* Lam. *Journal of Biological Chemistry* **261**: 8729–8733.

Vogt T, Jones P. 2000. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends in Plant Science* **5**: 380–386.

Wagner A, Ralph J, Akiyama T, Flint H, Phillips L, Torr K, Nanayakkara N, Te Kiri L. 2007. Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase in *Pinus radiata. Proceedings of the National Academy of Sciences, USA* **104**: 11856–11861.

Walker AM, Hayes RP, Youn B, Vermerris W, Sattler SE, Kang C. 2013. Elucidation of the structure and reaction mechanism of *Sorghum bicolor* hydroxycinnamoyltransferase and its structural relationship to other CoA-dependent transferases and synthases. *Plant Physiology* **162**: 640–651.

Wang JP, Naik PP, Chen HC, Shi R, Lin CY, Liu J, Shuford CM, Li Q, Sun YH, Tunlaya-Anukit S *et al.* 2014. Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in *Populus trichocarpa. Plant Cell* **26**: 894–914.

Weier D, Mittasch J, Strack D, Milkowski C. 2008. The genes *BnSCT1* and *BnSCT2* from *Brassica napus* encoding the final enzyme of sinapine biosynthesis: molecular characterization and suppression. *Planta* **227**: 375–385.

Weitzel C, Petersen M. 2011. Cloning and characterisation of rosmarinic acid synthase from *Melissa officinalis* L. *Phytochemistry* **72**: 572–578.

Withers S, Lu F, Kim H, Zhu Y, Ralph J, Wilkerson CG. 2012. Identification of grass-specific enzyme that acylates monolignols with *p*-coumarate. *Journal of Biological Chemistry* **287**: 8347–8355.

Yang Q, Reinhard K, Schiltz E, Matern U. 1997. Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA:anthranilate *N*-hydroxycinnamoyl/benzoyltransferase from elicited cell cultures of carnation, *Dianthus caryophyllus* L. *Plant Molecular Biology* **35**: 777–789.

Yang Q, Trinh HX, Imai S, Ishihara A, Zhang L, Nakayashiki H, Tosa Y, Mayama S. 2004. Analysis of the involvement of hydroxyanthranilate hydroxycinnamoyltransferase and caffeoyl-CoA 3-*O*-methyltransferase in phytoalexin biosynthesis in oat. *Molecular Plant–Microbe Interactions Journal* **17**: 81–89.

Yonekura-Sakakibara K, Tanaka Y, Fukuchi-Mizutani M, Fujiwara H, Fukui Y, Ashikari T, Murakami Y, Yamaguchi M, Kusumi T. 2000. Molecular and biochemical characterization of a novel hydroxycinnamoyl-CoA:anthocyanin 3-*O*-glucoside-6″-*O*-acyltransferase from *Perilla frutescens. Plant and Cell Physiology* **41**: 495–502.

Yu XH, Chen MH, Liu CJ. 2008. Nucleocytoplasmic-localized acyltransferases catalyze the malonylation of 7-*O*-glycosidic (iso)flavones in *Medicago truncatula. Plant Journal* **55**: 382–396.

Zhao J, Dixon RA. 2009a. The 'ins' and 'outs' of flavonoid transport. *Trends in Plant Science* **15**: 72–80.

Zhao J, Dixon RA. 2009b. MATE transporters facilitate vacuolar uptake of epicatechin 3′-*O*-glucoside for proanthocyanidin biosynthesis in *Medicago truncatula* and Arabidopsis. *Plant Cell* **21**: 2323–2340.

Zhao J, Huhman D, Shadle G, He XZ, Sumner LW, Tang Y, Dixon RA. 2011. MATE2 mediates vacuolar sequestration of flavonoid glycosides and glycoside malonates in *Medicago truncatula. Plant Cell* **23**: 1536–1555.

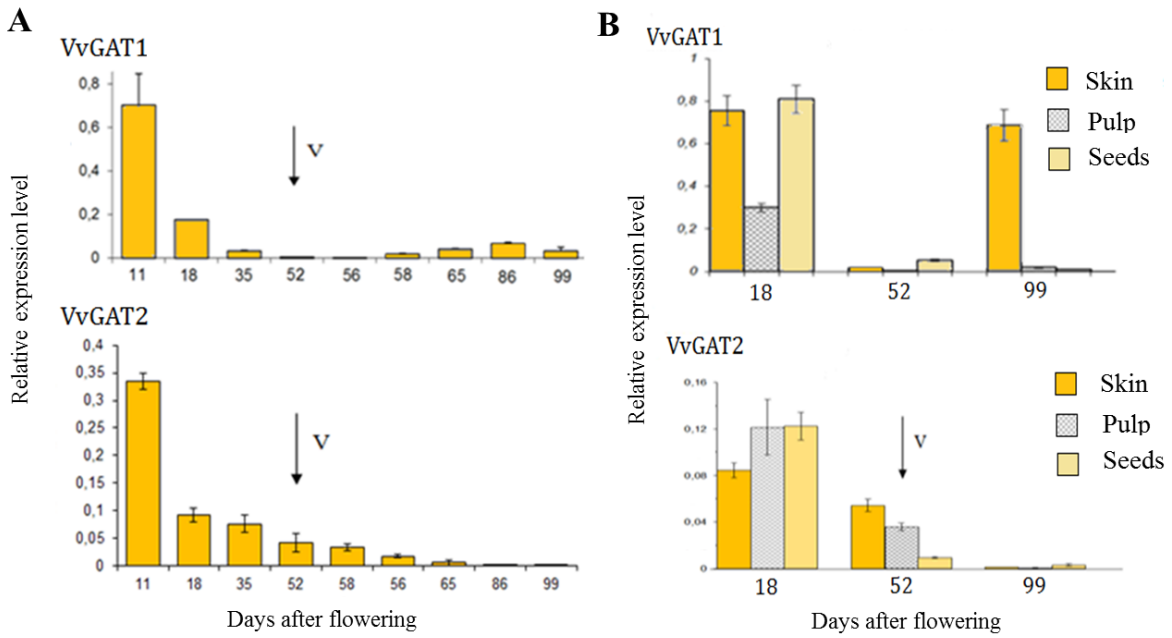


Figure 29. Expression pattern of *VvGAT1* and *-2*.
(**A**) berry pericarp along development, (**B**) skin, pulp and seeds during green stage (18 days after flowering, daf), at véraison (52 daf) and maturity (99 daf). The arrow indicates véraison.

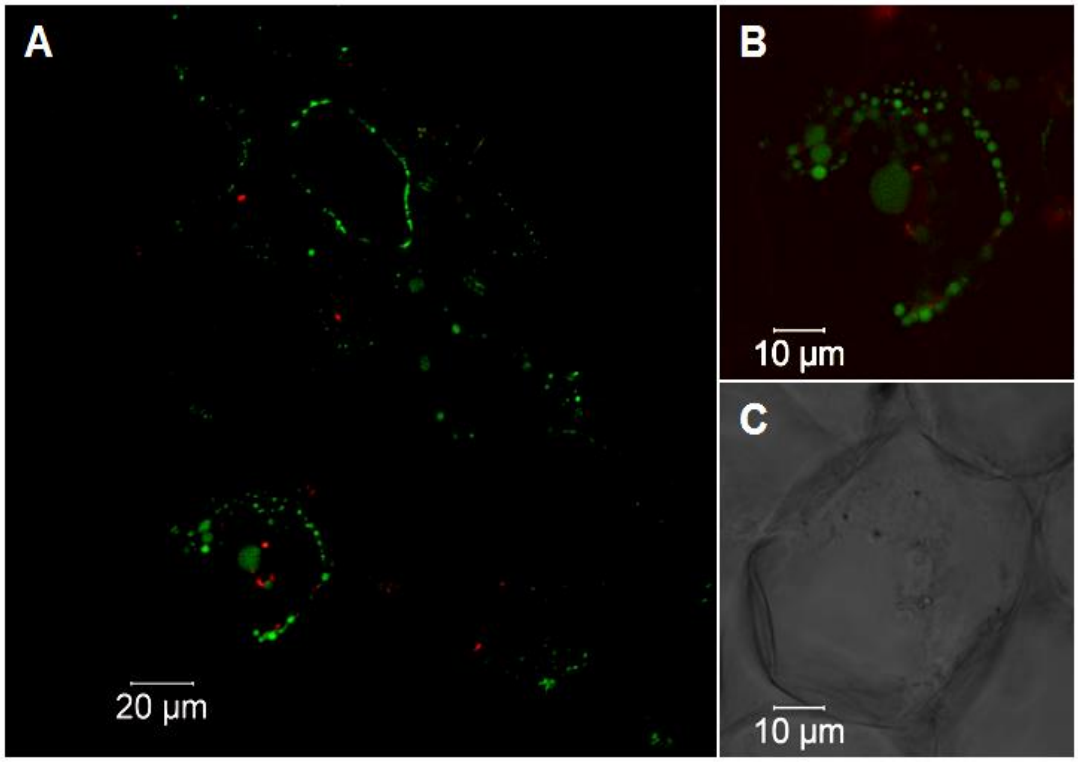


Figure 30. VvGAT1 localization in grapevine hairy-roots.
GFP fluorescence (green) and chlorophyll (red) were observed in grapevine hairy-roots transformed with VvGAT1-GFP construct. (**B**) is a zoom of (**A**). (**C**) corresponds to the same plan as (**B**) with transmitted light.

3. Previous results

The study of VvGATs started during a previous PhD work in the lab (Khater, 2009-2012).
Expression profile: In grape berry, *VvGATs* are mainly expressed during the green stage of pericarp development (11 daf) and their expression decreases until *véraison* (Fig. 29, A). Their expression stays very low after *véraison*.
VvGAT1 is mainly expressed in skin and seeds during the green stage (18 daf) and in skin at maturity (99 daf) (Fig. 29, B). *VvGAT2* is mainly expressed in pulp and seeds during the green stage (18 daf).

Intracellular localization: From hairy-roots transformed with the construction VvGAT1-GFP, VvGAT1 has been localized in cytosolic vesicles (0.5-0.9 µm), independent of chloroplasts (Fig. 30). It was hypothesized that the observed structures could be pre-vacuolar vesicles.

The heterologous expression of VvGATs in *Escherichia coli* and in the yeasts *Saccharomyces cerevisae* and *Pichia pastoris* did not lead to the production of functional proteins.
Several studies report post-traductional modifications for SCPL-ATs (glycosyation, pre-protein cleavage, disulfide bridge and heterodimer formation).
The fact that bacteria are devoid of cellular machinery able to ensure post-translational modifications could explain that the proteins expressed in *E.coli* were not active.
VvGAT production could be improved in yeast considering its codon preference and designing an optimized *VvGAT* sequence as monitored previously for an Arabidopsis sinapoyltransferase (Stehle *et al*., 2008).

In the present work, the expression of *VvGATs* has been performed in tobacco leaves (*Nicotiana benthamiana*). Indeed, a plant organism should be able to ensure suitable post-translational modifications. Another strategy to evaluate the impact of the encoded proteins *in planta* consisted in the overexpression of these genes in grapevine leaves.

The genes annotated as Serine Carboxypeptidases available in the 12X grapevine genome have been collected and aligned with already characterized SCPs and SCPL-ATs to construct a neighbor-joining tree. This phylogenetic study reveals that other grapevine genes could encode proteins functioning as SCPL-ATs.

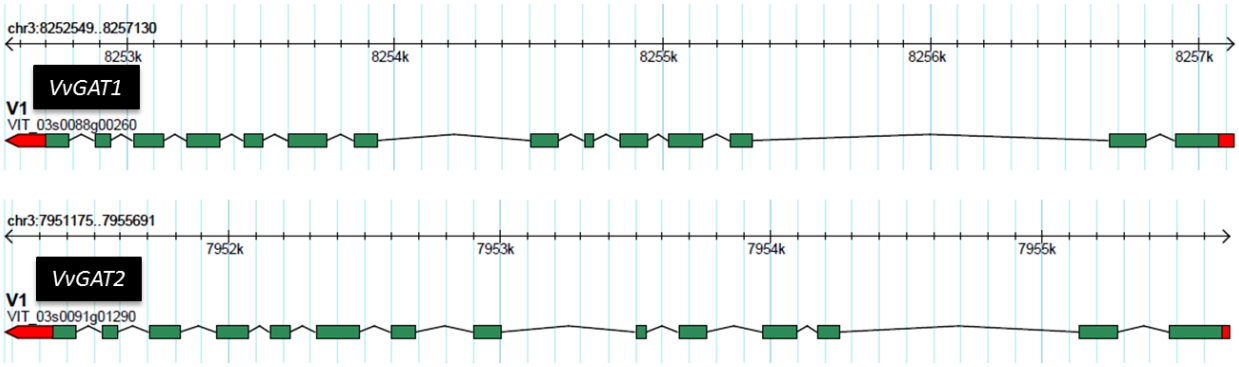


Figure 31. Genomic structure of VvGATs.

The exon-intron pattern was recovered from the 12X version of *Vitis vinifera* genome with gene prediction V1 (http://genomes.cribi.unipd.it/grape/). The scale indicates the chromosome and the genomic region shown.
Red and dark green boxes indicate the putative UTR regions and exonic sequences, respectively.

4. Research results

This part includes research results concerning Serine Carboxypeptidases which will be grouped with previous results in a manuscript for publication.

4. 1. Analysis of grapevine Serine Carboxypeptidases.

4. 1. 1. Focus on 2 serine carboxypeptidases.

Two genes annotated as serine carboxypeptidases were identified by transcriptomic screening by comparing grapevine hairy-roots overexpressing MybPA1 or -2 TFs (Terrier *et al*., 2009). Those TFs increased PA content in transgenic hairy-roots. This transcriptomic study was monitored with Qiagen Operon Array carrying only half of the complete set of grapevine genes. Both TFs induced genes related to the flavonoid pathway, notably *ANR* and *LAR1*, necessary to produce flavan-3-ols. VvGAT1 and -2 were the most induced genes annotated as serine carboxypeptidase. *VvGAT1* was induced by both MybPA1 and -2. Among the 306 genes significantly induced by MybPA1, *VvGAT1* was found in 44th position. *VvGAT1* was found as the 33rd most induced gene among the 159 genes significantly induced by MybPA2. *VvGAT2* was only induced by MybPA2 (39th).
A more recent transcriptomic analysis with Nimblegen microarray covering the whole grapevine genome revealed that *VvSDH8* and *-9*, *VvGTs* and *VvGATs* are induced by MybPA2 (Terrier, unpublished data). Interestingly, VvGAT1 and -2 still were the most induced genes annotated as serine carboxypeptidase.

The *VvGAT1* and *-2* genes correspond to VIT_03s0088g00260 and VIT_03s0091g01290, respectively, according to *Vitis Vinifera* 12X Genome Browser (http://genomes.cribi.unipd.it/grape/) and are located on chromosome 3.
VvGAT1 is located between positions 8,252,549 and 8,257,129 (4.58 kbp) and *VvGAT2* between positions 7,951,175 and_7,955,690 (4.515 Kbp). Each genomic sequence consists of 14 exons and 13 introns distributed differently (Fig. 31).
VvGATs were localized under a QTL influencing PAs galloylation in skin (Huang *et al*., 2012). However, the QTL interval covers a large part of the chromosome 3 that corresponds to hundreds of genes among which only a few genes or only one of them must be responsible for the observed phenotype.

An integrative approach, combining transcriptomics and quantitative genetics, allowed focusing on a restricted list of candidate genes for PAs biosynthesis (Carrier *et al*., 2013). *VvGATs* were included in this restricted list and *VvGAT1* was investigated by association genetics. This study revealed that the addition of a nucleotide (indel) in an intron was associated with the variation of PA galloylation.

VvGATs cDNAs were amplified from Macabeu cultivar (Mac). The cloned cDNA has a length of 1410 bp for VvGAT1 and 1467 bp for VvGAT2 encoding 470 and 489 amino acid long proteins, respectively. *In silico* analysis of protein sequences encoded by the cDNA predicts molecular masses of 53.13 kDa for VvGAT1 and 54.93 kDa for VvGAT2 (ExPASy server, Gasteiger *et al*., 2003).
VvGAT1 was also cloned from Portan (Prt) hairy-roots because PA %G is higher in Portan than in Maccabeu hairy-roots devoid of transgene (Dr Terrier, personal communication). A single difference

Table 14. Grapevine genes annotated as "serine carboxypeptidase" (12X version, annotation V1).

Name	VIT n°	Chromosome: Positions	Functional annotation	Protein length (aa)
VvGAT1	03s0088g00260	03: 08252549_08257129	Serine carboxypeptidase S10	470
VvGAT2	03s0091g01290	03: 07951175_07955690	Serine carboxypeptidase S10	489
VvSCP3	08s0040g01110	08: 12127838_12130203	Serine carboxypeptidase II	445
VvSCP4	08s0040g01060	08: 12093772_12096141	Serine carboxypeptidase S10	437
VvSCP5	03s0091g01240	03: 07903958_07907570	Serine carboxypeptidase S10	464
VvSCP6*	03s0088g00150	03: 08193838_08194272	Serine carboxypeptidase SCPL17	80
VvSCP7	01s0011g04900	01: 04557702_04559021	Serine carboxypeptidase-like 50 SCPL50	440
VvSCP8	13s0158g00070	13: 20955036_20961949	Serine carboxypeptidase	501
VvSCP9*	12s0034g02130	12: 18390381_18390989	Serine carboxypeptidase precursor	177
VvSCP10	18s0001g15550	18: 13667832_13671220	Serine carboxypeptidase II	468
VvSCP11*	03s0088g00160	03: 08195668_08197771	Serine carboxypeptidase SCPL17	223
VvSCP12*	11s0206g00110	11: 07448706_07453212	Serine carboxypeptidase S28	503
VvSCP13	08s0040g01140	08: 12150451_12157978	Serine carboxypeptidase K10B2.2 precursor	454
VvSCP14	11s0037g00800	11: 09327801_09351635	Serine carboxypeptidase II	471
VvSCP15	11s0037g00740	11: 09158796_09171285	Serine carboxypeptidase S10	501
VvSCP16	10s0042g01050	10: 14726735_14731533	Serine carboxypeptidase II	482
VvSCP17*	11s0037g00790	11: 09288160_09295060	Serine carboxypeptidase II	211
VvSCP18	06s0061g00520	06: 17985014_17987629	Serine carboxypeptidase precursor	501
VvSCP19	01s0011g04910	01: 04559022_04562444	Serine carboxypeptidase CPVL precursor	452
VvSCP20	13s0158g00080	13: 20963411_20967558	Serine carboxypeptidase	504
VvSCP21	01s0011g02000	01: 01710141_01715609	Serine carboxypeptidase S10	472
VvSCP22*	04s0023g03020	04: 19607568_19615885	Serine carboxypeptidase S28	488
VvSCP23*	13s0158g00060	13: 20952915_20953742	Serine carboxypeptidase SCPL40	65
VvSCP24	17s0000g00370	17: 00260913_00263712	Serine carboxypeptidase II	463
VvSCP25	07s0141g00520	07: 00296466_00302656	Serine carboxypeptidase 1 precursor	495
VvSCP26	14s0108g00110	14: 28916144_28920649	Serine carboxypeptidase S10	551
VvSCP27	03s0088g00050	03: 08035215_08039037	Serine carboxypeptidase 1	462
VvSCP28*	03s0088g00030	03: 08005949_08006486	Serine carboxypeptidase-like 6	496
VvSCP29	18s0001g11110	18: 09420352_09423762	Serine carboxypeptidase	488
VvSCP30*	11s0037g00780	11: 09317942_09323641	Serine carboxypeptidase II	392
VvSCP31	18s0001g08840	18: 07333925_07337286	Serine carboxypeptidase S10	469
VvSCP32	04s0023g00700	04: 17037859_17057049	Serine carboxypeptidase S10	535
VvSCP33	03s0063g00870	03: 04322365_04325281	Serine carboxypeptidase	488
VvSCP34	06s0061g00530	06: 17991854_17994476	Serine carboxypeptidase precursor	502
VvSCP35	11s0037g00320	11: 08125132_08131947	Serine carboxypeptidase S10	470
VvSCP36	05s0077g00560	05: 00372811_00377703	Serine carboxypeptidase S10	463
VvSCP37**	08s0040g01040	08: 12066763_12079925	Serine carboxypeptidase II	452
VvSCP38	13s0019g04040	13: 05307500_05309819	Serine carboxypeptidase 28	472
VvSCP39	13s0019g04030	13: 05301912_05306166	Serine carboxypeptidase SCPL27	457
VvSCP40	13s0158g00050	13: 20938716_20948513	Serine carboxypeptidase	497
VvSCP42	12s0055g00820	12: 14127382_14132040	Serine carboxypeptidase II	475
VvSCP43*	01s0010g00460	01: 15215062_15216231	Serine carboxypeptidase	355
VvSCP44	01s0011g02030	01: 01727310_01732878	Serine carboxypeptidase S10	477
VvSCP45	03s0088g00110	03: 08141417_08145242	Serine carboxypeptidase SCPL17	462
VvSCP46	13s0019g05130	13: 06884838_06892692	Serine carboxypeptidase III	449
VvSCP47		13: 06894900_06899920	Serine carboxypeptidase III	505
VvSCP48	08s0040g01070	08: 12098665_12104651	Serine carboxypeptidase S10	382
VvSCP50	14s0068g00260	14: 23972158_23975509	Serine carboxypeptidase S10	473
VvSCP51	03s0063g00860	03: 04315153_04317718	Serine carboxypeptidase	481
VvSCP52	11s0052g00770	11: 18321720_18324527	Serine carboxypeptidase SCPL13	461
VvSCP53	11s0065g00720	11: 14746919_14750209	Serine carboxypeptidase S10	456
VvSCP54	10s0116g00740	10: 00330532_00336864	Serine carboxypeptidase-like 32	501
VvSCP55	01s0244g00100	01: 21666026_21669055	Serine carboxypeptidase S10	464
VvSCP56	11s0052g00790	11: 18346277_18351149	Serine carboxypeptidase SCPL9	480
VvSCP57	11s0052g00750	11: 18301081_18305467	Serine carboxypeptidase 1 precursor	473
VvSCP58	14s0108g00100	14: 28913388_28915747	Serine carboxypeptidase S10	480
VvSCP59	08s0040g01130	08: 12137642_12140264	Serine carboxypeptidase II	449
VvSCP60	03s0091g01200	03: 07855565_07861870	Sinapoylglucose:malate sinapoyltransferase (SNG1)	492
VvSCP61	03s0091g01270	03: 07935398_07938767	Sinapoylglucose-choline *O*-sinapoyltransferase	468
VvSCP62	04s0023g00710	04: 17057671_17060682	Carboxypeptidase I precursor	483
VvSCP63	06s0061g00890	06: 18430619_18434504	Carboxypeptidase	459
VvSCP64*	08s0007g04480	08: 18402845_18412633	Pectinesterase family	821
VvSCP65*	13s0147g00120	13: 11134956_11150436	TIR-NBS-LRR-TIR disease resistance protein	982

*: Not included in phylogenetic study, *: Excluded after the first phylogenetic analysis, **: Re-annotated sequence included in phylogenetic analysis, between brackets: initial length of the re-annotated protein. Aa: amino acids.

has been detected between the VvGAT1 amino acid sequences of those 2 cultivars. VvGAT1_Prt carries a Proline at position 306 whereas VvGAT1_Mac carries a Serine. Grapevine genome has been sequenced from Pinot. VvGAT1 available in 12X version of grape genome also carries a Proline at position 306.

4. 1. 2. Phylogenetic analysis of grapevine serine carboxypeptidases

We searched all the genes annotated as serine carboxypeptidases (SCPs) according to *Vitis vinifera* 12X Genome Browser annotation V1 (http://genomes.cribi.unipd.it/grape/) and also by BLASTing the proteome, using VvGAT2 as a bait. Initially, 62 sequences annotated as SCPs have been found. The putative proteins encoded by these genes have been collected using "VIT" identifier in order to build a phylogenetic tree. The identity percentage between each sequence is presented in Annexes (Supplemental Figure 2).
Unlikely protein sequences resulting from gene annotation mistakes have been re-annotated and re-integrated into the phylogenetic analysis. A downstream start codon was identified in an exonic sequence for *VvSCP37* (VIT_08s0040g01040), which is likely the "real" start codon leading to a 452 amino acid long protein. VIT_13s0019g05130 has been separated in two sequences carrying 449 (VvSCP46) and 505 (VvSCP47) amino acids.
The serine carboxypeptidase domains of VvSCP64 and VvSCP65 are fused to a pectinesterase domain and TIR-NBS-LRR class disease resistance protein, respectively. Those two sequences could require further annotation and were not included in the phylogenetic analysis.
Other sequences leading to too short proteins could not be re-annotated and have not been included in the study (VvSCP6, -9, -11, -23, -30). VvSCP6 corresponds to the last 80 amino acids of a SCP protein but the beginning of the sequence is missing. This sequence shares 80% amino acid identity with VvSCP27 and -45. VvSCP9 is a 177 amino acid long sequence that can be aligned as 5 peptides on SCPs sequences. VvSCP9 exhibits 69% amino acid identity with VvSCP47. VvSCP11 corresponds to the beginning of a SCP sequence (223 amino acids). Its maximal identity is shared with VvSCP27 and -45 (69%). VvSCP23 is constituted of 65 amino acids and shares little homology (maximum 20%) with SCPs. VvSCP30 is shorter than usual SCPs (392 amino acids) but shares 51.4% homology with VvSCP14.
The first phylogenetic analysis revealed that some sequences exhibit high divergence compared to characterized SCPs (VvSCP12, -17, -22, -28, and -43). Indeed, their sequence exhibits less than 25% amino acids identity with the other SCPs and often lacks the characteristic GESYA (or GDSYS) motif, which corresponds to the catalytic center. These 5 sequences have been removed from the phylogenetic study.

Finally, 51 sequences annotated as SCP have been included in the phylogenetic analysis (Table 14). Fifty-one SCP genes were also detected in *Arabidopsis thaliana* genome (Fraser *et al*., 2005). The proteic sequences of already characterized plant SCPs and SCPL-ATs have been added to those of grapevine for phylogenetic analysis (Table 15). Not yet characterized SCPs from persimmon (DkSCPL1 and -2), putatively involved in flavan-3-ol galloylation, have been added in this analysis (Ikegami *et al*., 2007, Akagi *et al*., 2009). Otherwise, we included a SCP sequence from tea plant *Camellia sinensis* (CsSCPL, AIW39897) from NCBI database submitted by Chiu and coworkers (unpublished article entitled "Cloning and Molecular Characterization of a Novel Serine Carboxypeptidase-Like Gene in Oolong Tea Plant, *Camellia sinensis* (L.)").

Table 15. Characterized SCP proteins included in phylogenetic analysis.

Name	Species	NCBI n°	Function Or putative function	Reference
GAC	*Solanum pennellii*	AAF64227	Isobutyryl acylation	Li & Steffens, 2000
CtAT1	*Clitoria ternatea*	BAF99695	Anthocyanin acylation	Noda *et al*., 2007
SMT	*Arabidopsis thaliana*	AAF78760	Malate acylation	Lehfeldt *et al*., 2000 Hause *et al*., 2002
SCT	*Arabidopsis thaliana*	AAK52316	Choline acylation	Shirley & Chapple, 2003
BnSCT1	*Brassica napus*	AAQ91191	Choline acylation	Milkowski *et al*., 2004
BnSCT2	*Brassica napus*	CAM91991	Choline acylation	Weier *et al*., 2008
SAT	*Arabidopsis thaliana*	AEC07395	Anthocyanin acylation	Fraser *et al*., 2007
SST	*Arabidopsis thaliana*	AEC07397	Sinapoyl glucose acylation	
SCPL1	*Avena strigosa*		Avenacin acylation	Mugford *et al*., 2009, 2013
SCPL17	*Arabidopsis thaliana*	AAS99709	Glucosinolates acylation	Lee *et al*., 2012
DkSCPL1	*Diospyros kaki*	BAF56655	Flavan-3-ols galloylation	Ikegami *et al*., 2007
DkSCPL2	*Diospyros kaki*	BAH89272	Flavan-3-ols galloylation	Akagi *et al*., 2009
DgSCPL1	*Delphinium grandiflorum*	BAO 04182	Anthocyanin acylation	Nishizaki *et al*., 2013
DgSCPL2	*Delphinium grandiflorum*	BAO 04183	Anthocyanin acylation	
DgSCPL3	*Delphinium grandiflorum*	BAO 04184	?	
CsSCPL	*Camellia sinensis*	AIW39897	?	Chiu *et al*., unpublished
HvSCPI	*Hordeum vulgare*	P07519	Protease	Sørensen *et al*., 1986 Doan & Fincher, 1988
HvCPDWII	*Hordeum vulgare*	P55748	Protease	Sørensen *et al*., 1987 Liao & Remmington, 1990 Dal Degan *et al*., 1994
HvSCPIII	*Hordeum vulgare*	P21529	Protease	Sørensen *et al*., 1989
TaCPD-WII	*Triticum aestivum*	P08819	Protease	Breddam & Sørensen, 1987 Dóminguez *et al*., 2002
ScCPY	*Saccharomyces cerevisae*	NP_014026		Endrizzi *et al*., 1994 Sørensen & Breddam, 1997
SbHNL1	*Sorghum bicolor*	CAD12888	Hydroxynitrile lyase	Wajant *et al*., 1994
BRS1	*Arabidopsis thaliana*	Q9M099	Protease involved in brassinosteroids signaling	Li *et al*., 2001 Zhou & Li, 2005
SCPLe	*Solanum lycopersicum*	Q9M513	SCP Type I, wound inducible	Moura *et al*., 2001
MtSCP1	*Medicago truncatula*	XP_003592239	Mycorrhiza-specific SCP	Rech *et al*., 2013

In bold letters: Characterized SCPL-ATs.

Table 16. Putative serine carboxypeptidase clusters in grapevine genome.

Chromosome	Number of genes	Names	Cluster length(kb)
1	2	VvSCP7** and -19	4.742
1	2	VvSCP21 and -44	22.737
3	8	VvSCP60, VvSCP5, VvSCP61, GAT2, VvSCP27 VvSCP28*, VvSCP45, VvSCP6*, VvSCP11*, GAT1	401.564
3	2	VvSCP51 and -33	10.128
6	2	VvSCP18 and -34	9.462
8	6	VvSCP37**, -4, -48, -3, -59, -13	91.215
11	4	VvSCP15, -17*, -30*, -14	192.839
11	3	VvSCP57, -52, -56	50.068
13	2	VvSCP39 and -38	7.709
13	2	VvSCP46 and -47	14.93
13	5	VvSCP40, -23*, -8a**, -8b**, -20	28.842
14	2	VvSCP58 and -26	7.261

*: Partial sequence, not included in phylogenetic study, *: Excluded after the first analysis, **: Re-annotated sequence.

According to Fraser and coworkers (2005), 19 Arabidopsis SCPs clustered in a group called Clade Ia with tomato GAC, a SCPL-AT involved in the biosynthesis of different isobutyryl glucoses (Li & Steffens, 2000) and were identified as putative SCPL-ATs. Among them, they found already characterized Arabidopsis SCPL-ATs involved in the formation of sinapoyl choline (Shirley & Chapple, 2003), sinapoyl malate (Lehfeldt *et al*., 2000), whereas proteins not characterized as SCPL-ATs are excluded from this clade (Figure 32). In our study, all characterized SCPL-ATs, CsSCPL and DkSCPLs clustered in Clade Ia. Twelve VvSCPs clustered in this clade, including VvGAT1 and -2, arguing in favour of their putative acyltransferase function. VvGAT1 and VvGAT2 shared a maximum amino acid identity with CsSCPL from tea plant (64%) and DkSCPL1 from persimmon (61%), respectively. These two species produce galloylated flavan-3-ols, which would favour the hypothesis of the involvement of VvGATs in flavan-3-ol galloylation.

VvSCP31 is isolated on chromosome 18. It shares a maximal amino acids identity (57%) with CtAT1 which is involved in anthocyanin acylation with *p*-coumaroyl moiety in *Clitoria ternatea* (Noda *et al*., 2007). Otherwise, VvSCP31 is close to two SCPLs suspected to be involved in the transfer of hydroxybenzoic acid from its glucose ester to anthocyanins in *Delphinium grandiflorum* (Nishizaki *et al*., 2013). As a result, this putative grapevine SCPL could be involved in anthocyanin acylation. More precisely, it could encode a hydroxycinnamoyltransferase responsible for anthocyanin acylation. Indeed, anthocyanins can be acylated with caffeoyl or *p*-coumaroyl moieties in grapevine. Three VvSCPs clustering in Clade IA (VvSCP52, -56 and -57) share high amino acids identity between them (~70%) and are localized side by side in a 5 kb cluster on chromosome 11. Several QTLs influencing anthocyanin acylation have been recently highlighted in grapevine (Constantini *et al*., 2015). Notably, QTLs associated with anthocyanin coumaroylation have been detected on chromosome 3, 11 and 18. Further analysis is required to determine if the putative SCPL-ATs are responsible for this activity. *VvSCP5* expression is down-regulated in transgenic Shiraz berries in which *VvMybAs*, the anthocyanins specific TFs, are silenced (Rinaldo *et al*., 2015). Thus, the involvement of VvSCP5 in anthocyanin acylation is not excluded.

Two arabidopsis SCPs were grouped in Clade Ib (Fraser *et al*., 2005). VvSCP25, -32 and -62 clustered in Clade Ib with SCPI from *Hordeum vulgare* characterized as a peptidase and SCP from tomato (SCPLe). The last was identified as a type I SCP (Moura *et al*., 2001). Its expression was induced by wounding, systemin and methyl jasmonate. The encoded protein was localized in the central vacuole in leaves cells.

Seven VvSCPs clustered with carboxypeptidases from the yeast *Saccharomyces cerevisae* and SCPIII from *Hordeum vulgare*. We called this group Clade III.

Finally, the 29 other VvSCPs clustered in Clade II with CPDIII from *Hordeum vulgare* and *Triticum aestivum*, HNL1 from sorghum, BRS1 from Arabidopsis and MtSCP1 from barrel medic (*Medicago truncatula*). MtSCP1 is involved in arbuscular mycorrhiza symbiosis and must be secreted into the apoplast (Rech *et al*., 2013). SbHNL1 is a hydroxynitrile lyase involved in the cleavage of cyanohydrins and thus in the release of hydrogen cyanide (HCN), a compound used as a defense against herbivores. To sum up, none of the proteins clustered in Clade Ib, II and III have been characterized as acyltransferase.

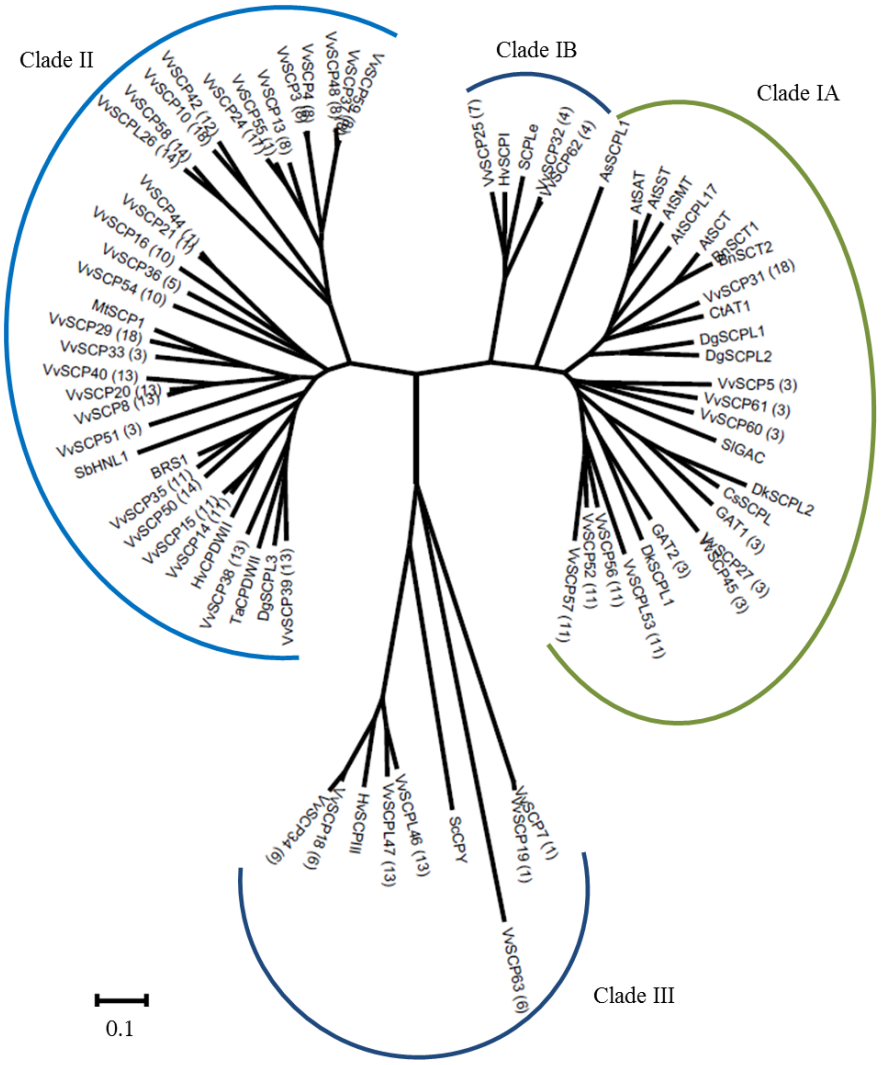


Figure 32. Serine carboxypeptidases neighbor-joining tree.

For grapevine SCPs, the chromosome is indicated between brackets. The scale bar represents 0.1 substitutions per site.

Abbreviations are: VvGAT1: VIT_03s0088g00260; VvGAT2: VIT_03s0091g01290; VvSCP3: VIT_08s0040g01110; VvSCP4: VIT_08s0040g01060; VvSCP5: VIT_03s0091g01240; VvSCP7: VIT_01s0011g04900; VvSCP8: VIT_13s0158g00070; VvSCP10: VIT_18s0001g15550; VvSCP13: VIT_08s0040g01140; VvSCP14: VIT_11s0037g00800; VvSCP15: VIT_11s0037g00740; VvSCP16: VIT_10s0042g01050; VvSCP18: VIT_06s0061g00520; VvSCP21: VIT_01s0011g02000; VvSCP24: VIT_17s0000g00370; VvSCP25: VIT_07s0141g00520; VvSCP26: VIT_14s0108g00110; VvSCP27: VIT_03s0088g00050; VvSCP29: VIT_18s0001g11100; VvSCP31: VIT_18s0001g08840; VvSCP32: VIT_04s0023g00700; VvSCP33: VIT_03s0063g00870; VvSCP34: VIT_06s0061g00530; VvSCP35: VIT_11s0037g00320; VvSCP36: VIT_05s0077g00560; VvSCP37: VIT_08s0040g01040; VvSCP38: VIT_13s0019g04040; VvSCP39: VIT_13s0019g04030; VvSCP40: VIT_13s0158g00050; VvSCP42: VIT_12s0055g00820; VvSCP44: VIT_01s0011g02030; VvSCP45: VIT_03s0088g00110; VvSCP46 and VvSCP47: VIT_13s0019g05130; VvSCP48: VIT_08s0040g01070; VvSCP50: VIT_14s0068g00260; VvSCP51: VIT_03s0063g00860; VvSCP52: VIT_11s0052g00770; VvSCP53: VIT_11s0065g00720; VvSCP54: VIT_10s0116g00740; VvSCP55: VIT_01s0244g00100; VvSCP56: VIT_11s0052g00790; VvSCP57: VIT_11s0052g00750; VvSCP58: VIT_14s0108g00100; VvSCP59: VIT_08s0040g01130; VvSCP60: VIT_03s0091g01200; VvSCP61: VIT_03s0091g01270. SlGAC: *Solanum lycopersicum* glucose acyltransferase (AAF64227); CtAT1: *Clitoria ternalea* Acyltransferase 1(BAF99695, Patent Publication Number WO/2007/046148); AtSMT: *Arabidopsis thaliana* 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase (AAF78760); AtSCT: *Arabidopsis thaliana* 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase (AAK52316); BnSCT1: *Brassica napus* 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase 1 (AAQ91191); BnSCT2 (CAM91991); AtSAT: *Arabidopsis thaliana* 1-*O*-sinapoyl-β -glucose:anthocyanin sinapoyltransferase (AEC07395); AtSST: *Arabidopsis thaliana* 1-*O*-sinapoyl-β -glucose:1-*O*-sinapoyl-β-glucose sinapoyltransferase (AEC07397); AsSCPL1: *Avena sativa* Saponin-deficient 7 (ACT21078); DkSCPL1: *Diospyros kaki* Serine Carboxypeptidase Like 1 (BAF56655); DkSCPL2 (BAH89272); AtSCPL17 (AAS99709); CsSCPL: *Camellia sinensis* SCPL (AIW39897); DgSCPL1: *Delphinium grandiflorum* Serine Carboxypeptidase Like 1 (BAO 04182); DgSCPL2 (BAO 04183); DgSCPL3 (BAO04184). HvSCPI: *Hordeum vulgare* serine carboxypeptidase I (P07519); HvCPDWII: *Hordeum vulgare* serine carboxypeptidase II (P55748); HvSCPIII: *Hordeum vulgare* serine carboxypeptidase III (P21529); TaCPD-WII: *Triticum aestivum* serine carboxypeptidase II (P08819); ScCPY: *Saccharomyces cerevisae* carboxypeptidase Y (NP_014026); SbHNL1: *Sorghum bicolor* Hydroxynitrile lyase (CAD12888); BRS1: *Arabidopsis thaliana* brassinosteroid-insensitive 1 (Q9M099); SCPLe: serine carboxypeptidase *Lycopersicum esculentum* (Q9M513); MtSCP1: *Medicago truncatula* serine carboxypeptidase 1 (XP_003592239)

4. 2. 3. Genomic analysis

VvSCPs were located on 14 chromosomes.
Independently of clades, we observed that some SCPs located on the same chromosome are physically grouped (Table 16). This could result from gene duplication yielding gene clusters. On chromosome 3, 10 putative SCPL-ATs (including VvSCP6, -11 and -28, not included in the neighbor-joinig tree) were found in ~400 kb (Annexes, Supplemental Figure 3). VvSCP28 shared little amino acids identity with the other SCPs and is unlikely to have a SCP activity. VvSCP6 and -11 are partial sequences. The other 7 sequences share about 50% sequence identity, even if VvSCP27 and -45 share 98.9% sequence identity. Only one amino acid substitution has been identified between these two sequences.
Another cluster was localized in chromosome 11, with 3 putative SCPL-ATs grouped in ~50 kb. VvSCP52, -56 and -57 share between 63.5 and 69.1 % of proteic sequence identity.

4. 2. 4. Comparison of key amino acids

The protein sequences used for phylogenetic analysis were aligned in order to compare the amino acids necessary for the activity of SCPs.
SCPs and SCPL-ATs share a common ancestor gene and SCPL-ATs have acquired the capacity to catalyse acyltransferase activity instead of hydrolysis. The molecular bases involved in this specification have been addressed and investigated (Stehle *et al*., 2006). Both SCPs and SCPL-ATs possess a catalytic triad of 3 non successive amino acids Ser-Asp-His. Site-directed mutagenesis studies reveal that Ser and His residues are crucial for acyltransferase activity whereas the activity is less affected by Asp mutagenesis (Stehle *et al*., 2006). The catalytic triad is well conserved among all the analysed sequences except for DkSCPL1 which is devoid of the 3 amino acids which usually constitute the catalytic triad, VvGAT2 which is devoid of the His residue and VvSCP29 which is devoid of the Ser residue. However, VvGAT2 carry a His residue 7 amino acids further in its sequence, contrary to other putative SCPL-ATs (except DkSCPL2 which possess a His residue 6 amino acids further).
VvGAT1 possess the complete catalytic triad. The mechanism described by Stehle and coworkers (2009) for malate sinapoylation by AtSMT could be transposed to epicatechin galloylation from β-G (Fig. 33).
The substitution of Glu by Asp in the catalytic centre, found in characterized SCPL-ATs has been proposed as a key modification involved in glucose ester recognition (Stehle *et al*., 2006). The sequences found in clade Ia exhibit a Glu instead of Asp in the catalytic centre, whereas members of the 3 other clades exhibit an Asp (Fig. 34).

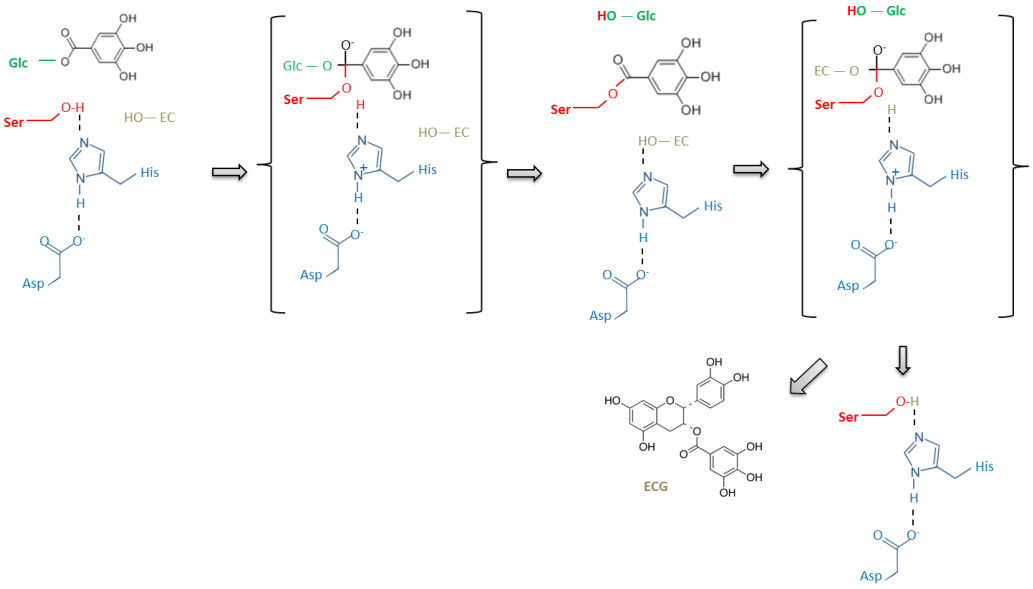


Figure 33. Scheme of the hypothetic mechanisms occurring during SCPL-AT catalysed galloylation of flavan-3-ols.
This figure is adapted from Stehle and coworkers (2009). Once activated by Histidine (in blue), the seryl residue (in red) attacks the acyl donor (β-glucogallin) to form a tetrahedral state (between curly brackets). Glucose (in green) is released by cleavage and Serine is bound to galloyl moiety (acyl enzyme intermediate). The acyl acceptor (EC, epicatechin, in brown) is activated by Histidine and attacks the acyl enzyme intermediate to form a second tetrahedral state. Finally, the seryl residue leaves the active site to return to its initial position, and reaction product (ECG, epicatechin gallate) is released.

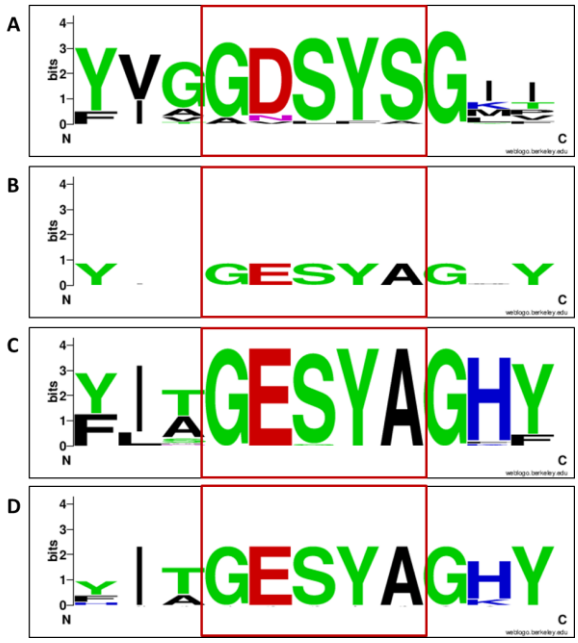


Figure 34. Consensus motifs for each of the major clades generated with Weblogo.
A. Clade Ia. B. Clade Ib. C. Clade II. D. Clade III. SCPL-AT catalytic center is framed. The Ser found in the middle of the catalytic center belongs to the catalytic triad. Each logo consists of stacks of letters, one stack for each position in the sequence (Crooks *et al.*, 2004). The overall height of each stack indicates the sequence conservation at that position (measured in bits), whereas the height of symbols within the stack reflects the relative frequency of the corresponding amino acid at that position. Amino acids have colors according to their chemical properties (Lewin 1994); polar amino acids (G, S, T, Y, C, Q, N) show as green, basic (K, R, H) blue, acidic (D, E) red, and hydrophobic (A,V, L, I, P, W, F, M) amino acids as black.

The predicted signal peptide includes the first 33 and 17 amino acids of VvGAT1 and -2, respectively (Annexes, Supplemental Table 6).
VvGAT1 and -2 carry 3 and 4 potential glycosylation sites (Asn-X-Ser/Thr), respectively (Annexes, Supplemental Figure 4). However, these sites are not located in the same positions.
Concerning VvGAT1, the first site includes amino acids 45 (Asn), 46 (Leu) and 47 (Thr) according to VvGAT1 sequence numerotation. No other putative SCPL-AT possesses a glycosylation site at this position, but a glycosylation site was found for VvSCP14, -15, -26 and -58 (clade II) at this position.
The second site is located at amino acids 356-358 and is shared with 12 other SCPs from clade IA.
The last potential glycosylation site (406-408) is shared by 17 SCPs from group IA.
Concerning VvGAT2, the first glycosylation site included amino acids 137, 138 and 139, according to VvGAT2 sequence numerotation. DkSCPL1 and AsSCPL1 also carry this glycosylation site.
The 2nd glycosylation site (250-252) is shared by DkSCPL1 and -2.
The 3rd glycosylation site (297-299) is shared with VvSCP5 only.
The last site (368-370) is also shared by 13 other SCPs from clade IA (notably SlGAC, AtSAT and AtSMT) and 3 members of clade IB.
In summary, 3 potential glycosylation sites are well conserved among putative SCPL-ATs (clade IA).

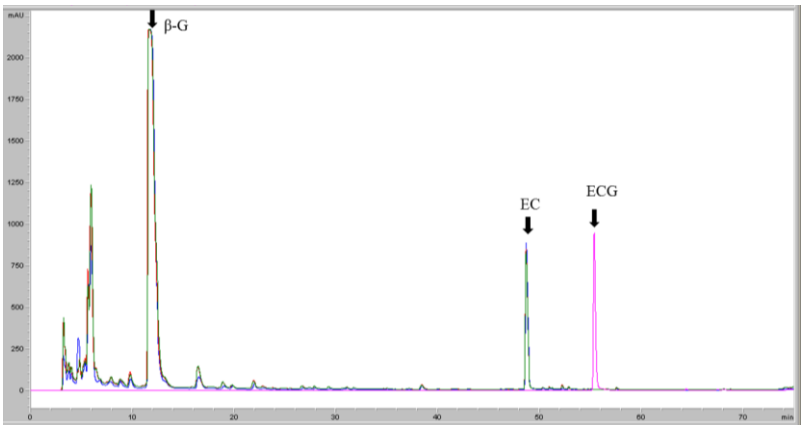
4. 2. Functional validation of VvGATs.

4. 2. 1. Transient transformation of tobacco leaves.

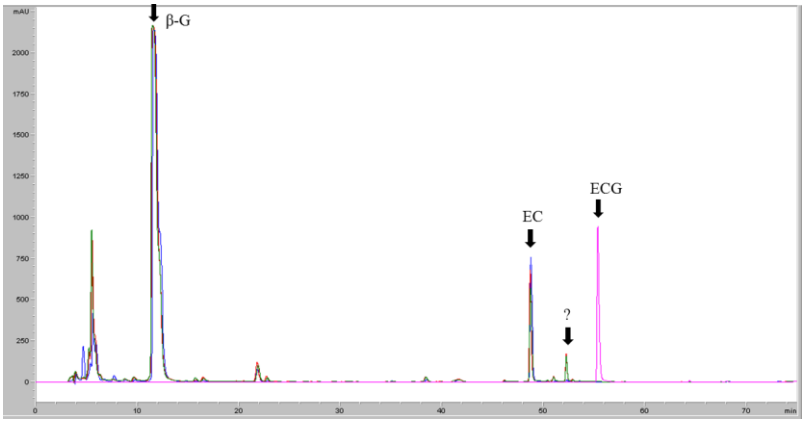
4. 2. 1. 1. *In vitro* assays with proteins extracts from inoculated tobacco leaves

VvGAT1 and *-2* were cloned from Macabeu cDNA in pEAQ-HT-DEST1, an optimized vector for transient expression of proteins in tobacco (Sainsbury *et al*., 2009).
Each vector has been incorporated in *Agrobacterium tumefaciens* GV3101. The agrobacteria solution was infiltrated at the underside of 4-6 weeks old tobacco leaves. 5-7 days after inoculation, the inoculated areas were harvested, weighed, frozen in liquid nitrogen, ground to a fine powder and stored at -80°C until analysis. Total proteins were extracted, desalted and quantified according to the Bradford method.
Galloylation reaction was tested in presence of β-G and EC as substrates. The method used for each assay is presented in Material and methods (Chapter 2).
The reaction products were analysed by HPLC. Three independant galloylation assays are presented in Fig. 35 (A-C).

A



B



C

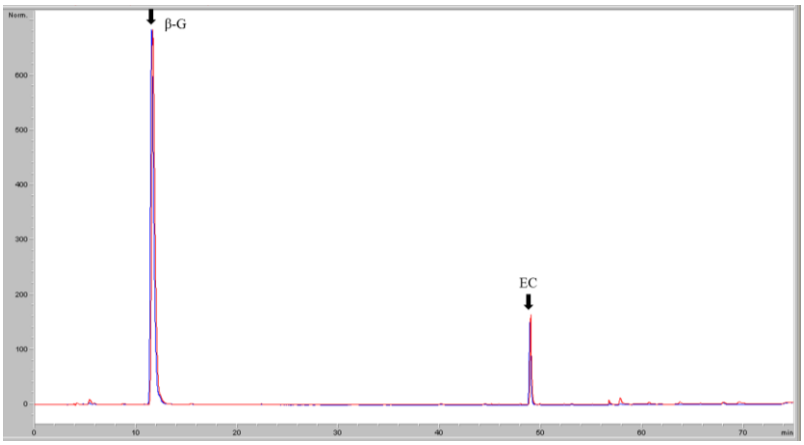


Figure 35. HPLC-DAD analysis of epicatechin galloylation assay from tobacco leaves proteic extracts.
Enzymatic assay was tested with proteins extracted from tobacco leaves inoculated with pEAQ-HT-GFP (control, blue line), pEAQ-HT-DEST1:GAT1 (red line) or pEAQ-HT-DEST1:GAT2 (green line). A solution of epicatechin gallate (ECG) was also injected to identify its retention time (pink line).The reaction mixture contained 1 mM β-glucogallin (β-G), 0.4 mM epicatechin (EC) and 4 mM ascorbic acid. After 2h incubation at 30°C, the reaction was stopped by addition of acetone:water:TFA (70:30:0.01, v/v/v) and the solution evaporated to dryness with Genevac (except condition **C**). The pellet was conserved at -20°C until analysis. The resuspended pellets in methanol:water:TFA (20:80:0.05, v/v/v) were injected on HPLC-DAD system and analysed at 280 nm.**A**. The reaction was tested in potassium phosphate buffer (50 mM, pH 6) with about 400 µg of desalted proteins (PD Mini-Trap G25 column). **B**. The reaction was tested in MES-KOH buffer (50mM, pH 5), started with about 50-200 µg of desalted proteins (PD Mini-Trap G25 column). ? : unknown compound further subject to UPLC-DAD-MS analysis. **C**. The reaction was tested in MES-KOH buffer (50 mM, pH 5.6) with 15-20 µg of desalted proteins (Amicon Ultra column, 10 kDa, Merck Millipore).

No peak with a retention time corresponding to ECG has been identified.
A peak corresponding to an unknown compound has been detected in the chromatograms corresponding to the condition containing enzyme extract from tobacco leaves overexpressing VvGAT1 and VvGAT2 (Fig. 35, B). The reactions with protein extracted from tobacco leaves inoculated with pEAQ-HT-GFP or pEAQ-HT-DEST1:GAT1 have been further injected in UPLC-DAD-MS system to characterize this unknown peak (Annexes).

The UPLC-DAD-MS analysis of the unknown compound revealed that this compound is also found in control condition (leaves inoculated with pEAQ-HT-GFP vector) but in lower amount, according to peak area (Annexes, Supplemental Figure 5). The fragmentation pattern revealed that this molecule could be composed of a glucosyl part, but the identity of the molecule could not be determined.

The analysis of the chromatograms resulting from *in vitro* assay revealed that in none of the tested conditions, ECG was produced (Fig. 35).
Those results led us to formulate the hypothesis that VvGAT was not produced, produced but not active, or that *in vitro* reaction conditions were not suitable.
To check these hypotheses, we monitored transcriptomic and proteomic analysis from tobacco leaves extracts.

4. 2. 1. 2. *VvGAT* expression in tobacco leaves

The relative expression level of *VvGAT* in transformed tobacco leaves has been checked by qPCR using *Protein phosphatase 2A* (*PP2A*) as reference gene (Liu *et al*., 2012, Fig. 36). Non-inoculated leaves (NI) and leaves from 2 independent tobacco plants inoculated with pEAQ-HT-GFP vector (GFP1 and 2) have been used as controls.
VvGATs relative expression has been tested with 1 or 2 inoculated leaves of 3 independent plants.
VvGATs relative expression was null in control leaves whereas each gene was well expressed in tobacco leaves transformed with its corresponding pEAQ-HT-DEST1:VvGAT construct. We observed variability among the samples probably due to infiltration efficiency.

4. 2 .1. 3. Analysis of total proteins from tobacco leaves

The proteins extracts from non-inoculated leaves and leaves inoculated with pEAQ-HT-DEST1: GAT1 or -2 was desalted using an adequate column (PD Mini-Trap G25 column, GE Healthcare Life Sciences). After desalting, the most concentrated proteins extracts (3 aliquots) were loaded on SDS-Page gel (Fig. 37). No additional band that could correspond to VvGAT proteins has been detected in transiently transformed tobacco leaves.

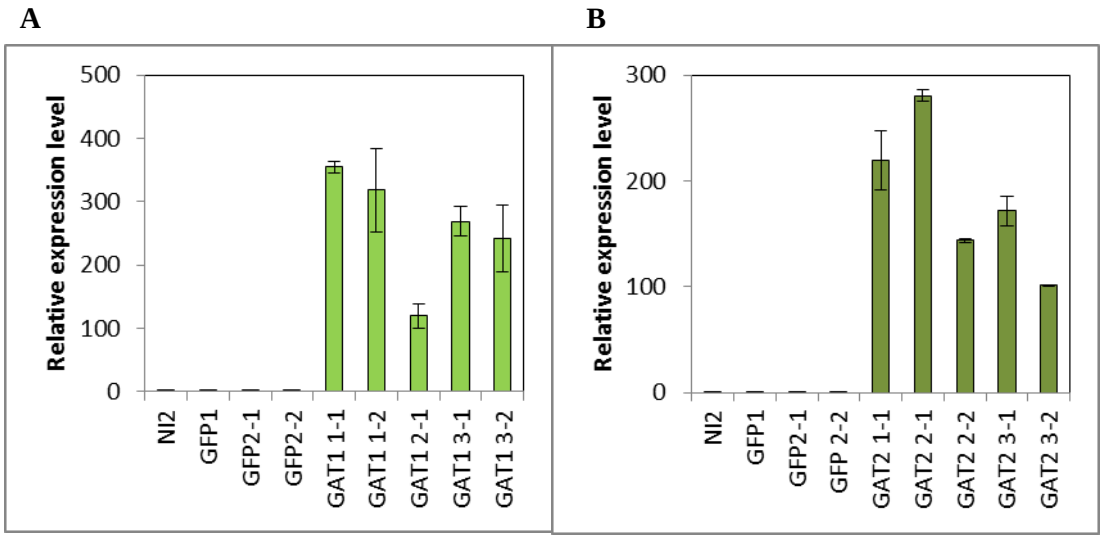


Figure 36. Relative expression level of VvGATs in tobacco leaves.
NI: Non-inoculated. GFP: Leaves inoculated with agrobacteria carrying pEAQ-HT-GFP vector. **A**. GAT1: Leaves inoculated with agrobacteria carrying pEAQ-HT-DEST1:VvGAT1 construct. **B**. GAT2: Leaves inoculated with agrobacteria carrying pEAQ-HT-DEST1:VvGAT2 construct.

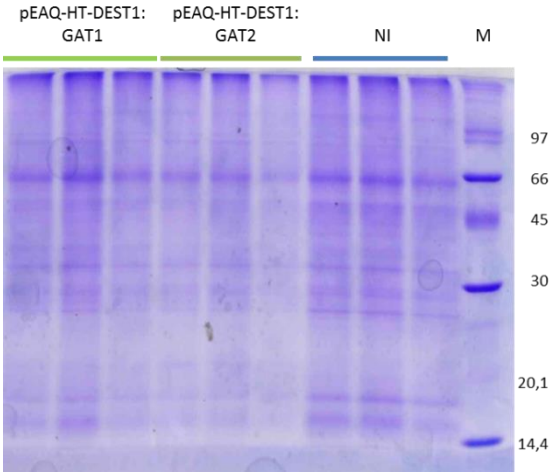


Figure 37. Analysis of total proteins extracted from tobacco leaves with SDS-PAGE.
The 3 most concentrated protein extracts, according to Bradford assay, have been deposited on 3 independent wells. pEAQ-HT-DEST1 GAT1: Tobacco leaves infiltrated with agrobacteria carrying the construct pEAQ-HT-DEST1 GAT1, pEAQ-HT-DEST1 GAT2: Tobacco leaves infiltrated with agrobacteria carrying the construct pEAQ-HT-DEST1 GAT2, NI: Non inoculated tobacco leaves, M: Ladder.

4. 2. 2. Identification of phenolic compounds in tobacco leaves.

Due to the lack of activity from tobacco proteins extracts, we tried to measure the effect of VvGATs on endogenous phenolic compounds in tobacco leaves.
Firstly, we identified the main phenolic compounds in tobacco leaves.
Non inoculated leaves extract were prepared for UPLC-DAD-MS analysis. Fourteen peaks were identified at 278nm using UPLC-DAD-MS system (Fig. 38). 6 peaks (n° 1, 2, 6, 8, 10 and 12) had a $[M-H]^-$ = 353 that corresponds to caffeoyl quinic acids.
The presence of other hydroxycinnamoyl quinic acids has been described in *Nicotiana benthamiana* leaves (Ncube *et al*., 2014). The presence of *p*-coumaroyl and feruloyl quinic acids has been screened by the research of signals corresponding to $[M-H]^-$=337 and 367, respectively. None of those ions has been detected in the leaf extracts.
Different forms of CQAs have been described in tobacco, depending on the carbon number (3, 4 or 5) whose –OH group is acylated, and the caffeoyl moiety configuration (*cis* or *trans*). According to the elution order (with C18 column) and the fragmentation pattern of those compounds (Clifford *et al*., 2005), we identified 5 CQAs (Table 16).
trans-4-caffeoylquinic acid (trans-4-CQA, peak 10) and *trans*-5-caffeoylquinic acid (*trans*-5-CQA, peak 8) exhibit the most important peak area.

4. 2. 3. Analysis of transformed tobacco leaves.

Once the main CGAs identified, we measured the effect of VvGAT overexpression on those phenolic compounds *in planta*. The content of endogenous CQAs has been compared between tobacco leaves inoculated with the construction pEAQ-HT-GFP (control), pEAQ-HT-DEST1:GAT1 or GAT2. This experiment was monitored to test if VvGATs could increase CQAs content *in planta*, by transfer of a caffeoyl group from caffeoyl glucose ester to quinic acid.
The 4 major CQAs have been quantified: *trans*-3-CQA, 4-CQA (*trans* and *cis* isomers) and *trans*-5-CQA (Fig. 39). VvGAT1 and VvGAT2 tended to increase the content of each quantified CQAs. However, we observed a large heterogeneity between individuals and the increase was never significant.

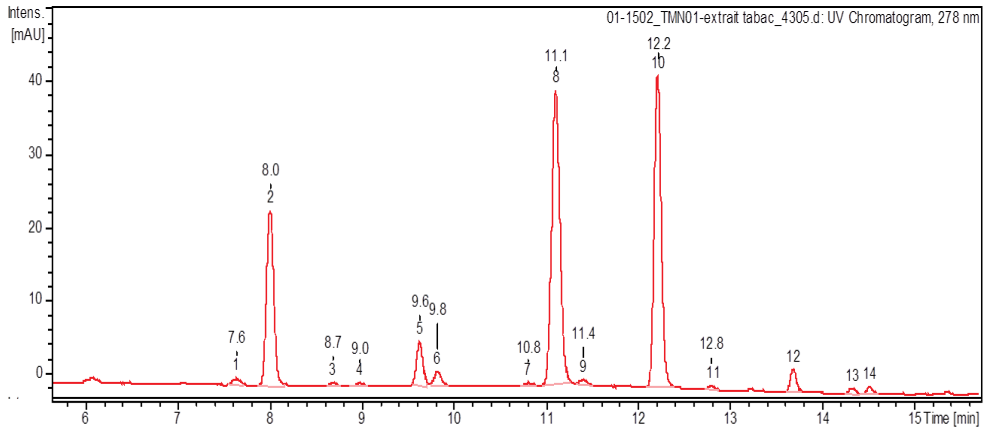


Figure 38. UPLC-DAD-MS analysis of phenolic compounds from non inoculated tobacco leaves.
The red line corresponds to the chromatograms at 278 nm.

Table 16. Characterization of chlorogenic acids detected in tobacco leaves by UPLC-DAD-MS analysis according to fragmentation spectra.

Molecule	**Peak n°**	**Retention time**	**[M-H]$^-$**	**m/z of fragment ions (% intensity)**
cis-3-CQA	1	7.6	353	191 (100%)
				179 (78 %)
trans-3-CQA	2	8	353	191 (100%)
				179 (64 %)
cis-4-CQA	6	9.8	353	173 (100%)
				179 (50%)
				191 (19%)
trans-4-CQA	10	12.2	353	173 (100%)
				179 (50%)
				191 (12%)
trans-5-CQA	8	11.1	353	191 (100%)
cis-5-CQA	12	13.7	353	191 (100%)
				179 (16%)

CQA : caffeoyl quinic acid.

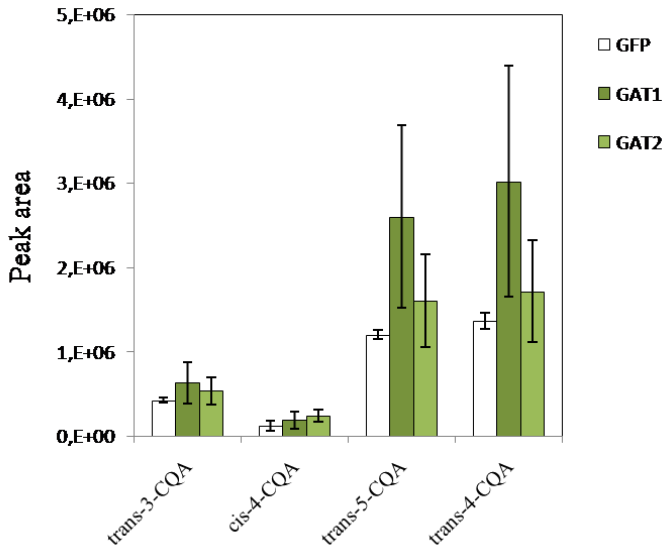


Figure 39. Caffeoylquinic acids content in tobacco leaves analysed with HPLC-DAD system.
Leaves were infiltrated with a solution of agrobacteria carrying pEAQ-HT-GFP vector (GFP, white), pEAQ-HT-DEST1:GAT1 (GAT1, dark green), or pEAQ-HT-DEST1:GAT2 (GAT2, light green). Peaks areas were quantified at 280 nm. Each bar corresponds to the mean value of 3 biological replicates ± standard deviation.

4. 2. 4. Substrates infiltration in inoculated tobacco leaves.

Tobacco leaves inoculated with the previously described constructs have been infiltrated 3 days after inoculation with potential substrates of VvGATs (β-G and EC). Then, the just infiltrated leaves have been cut and their petiole has been dipped in the substrates mixture overnight.
Thus, we provided substrates for VvGATs and expected an enzymatic reaction *in planta*. This method is inspired from Karamat *et al*. (2014) to validate the function of a prenyltransferase.
The presence of the infiltrated substrates (β-G and EC) and the expected product (ECG) has been monitored by extracted-ion chromatograms (EIC) at the m/z values corresponding to each compound (i.e. 331, 289, and 441, respectively for β-G, EC, and ECG, in the negative ion mode) for the different conditions.
As expected, EC and β-G have been found in the leaves infiltrated and dipped in substrates mixture (Fig. 40 and 41, A, B and D) but were absent in leaves not infiltrated with this mixture. It confirmed that the putative substrates of VvGATs are not naturally produced in tobacco leaves and that their infiltration was necessary to test galloylation activity *in planta*.
An ion that exhibits the same molecular mass as epicatechin 3'-*O*-glucoside, previously identified in *Medicago truncatula* (Pang *et al*., 2008), has been detected in extracts from leaves infiltrated with the substrates mixture at m/z 451 in the negative ion mode (Annexes, Supplemental Figure 6). However, the corresponding peak was coeluted with a peak of CGA and we could not check its UV absorption spectrum.
EIC monitored for ECG did not yield a peak corresponding to this flavan-3-ol (data not shown).

4. 2. 5. Transient expression of *VvGATs* in grapevine leaves.

As tobacco leaves infiltration did not allow functional characterization of VvGATs, we have tested agroinfiltration in grapevine leaves, following a method developed by Santos-Rosa *et al*. (2008).
Phenolic compounds content has been compared between control leaves (transformed with pEAQ-HT-GFP) and leaves transformed with a gene of interest.
Two independent experiments were monitored.
In the first experiment, only one analysis could be performed due to the lack of plant material. Indeed, the two younger leaves of a grapevine plantlet provide only 25-30 mg of fresh weight and 20mg of frozen leaf powder were necessary for extraction of phenolic compounds (data not shown). The leaves transformed with *VvGAT1_Mac* contained less PAs but with a higher %G compared to the leaves transformed with pEAQ-HT-GFP (control condition, Table 17). On the contrary, they accumulated less hydroxycinnamoyl tartrates (about half) than control leaves. We cannot exclude a problem during the extraction procedure since we observed a lower amount of phenolic material extracted from leaves transformed with *VvGAT1_Mac*.

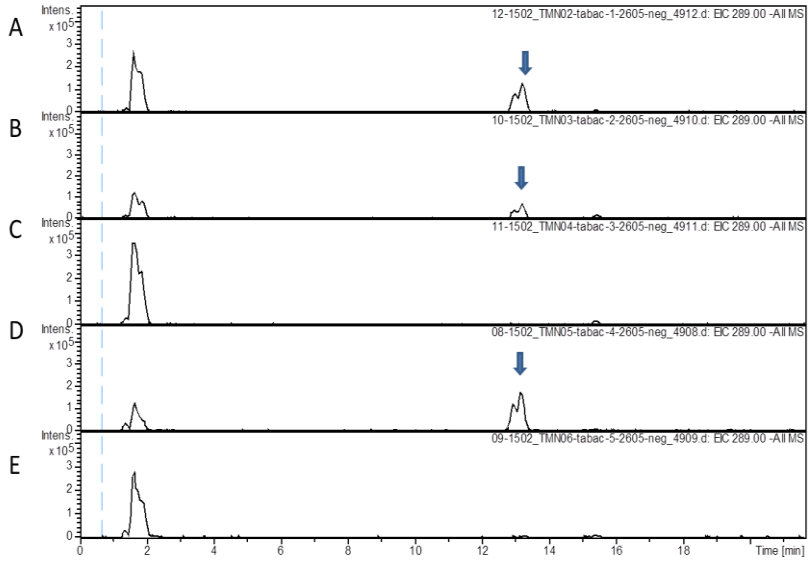


Figure 40. Extracted-ion chromatogram (EIC) for (-)-epicatechin (at m/z= 289 a.m.u) in tobacco leaves.
The arrow shows the peak corresponding to (-)-epicatechin. In conditions **A**, **B** and **D**, a mixture of (-)-epicatechin and β-glucogallin (final concentration 200 µM) has been infiltrated 3 days after inoculation in the same inoculated leaves zones. After substrates infiltration, the leaves have been detached and their petiole dipped into the mixture. **A**. Non inoculated leaves infiltrated and dipped in the substrates mixture. **B**. Leaves inoculated with pEAQ-HT-GFP vector and infiltrated and dipped in the substrates mixture. **C**. Leaves inoculated with pEAQ-HT-GFP vector. **D**. Leaves inoculated with pEAQ-HT-DEST1:GAT1 construct and infiltrated and dipped in the substrates mixture. **E**. Leaves inoculated with pEAQ-HT-DEST1:GAT1 construct.

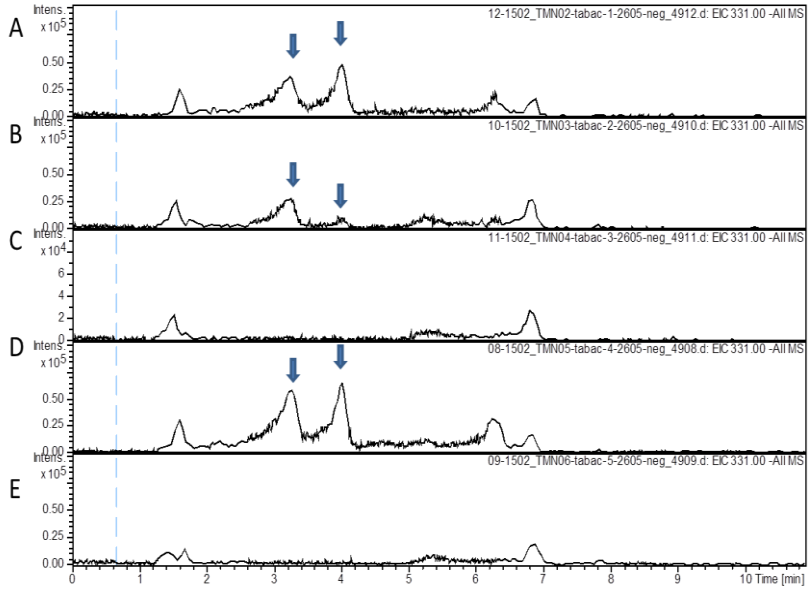


Figure 41. Extracted-ion chromatogram for β-glucogallin (at m/z= 441 a.m.u) in tobacco leaves.
The arrows show the peaks corresponding to β-glucogallin.
In conditions **A**, **B** and **D**, a mixture of (-)-epicatechin and β-glucogallin (final concentration 200 µM) has been infiltrated 3 days after inoculation in the same inoculated leaves zones. After substrates infiltration, the leaves have been detached and their petiole dipped into the mixture.
A. Non inoculated leaves infiltrated and dipped in the substrates mixture. **B**. Leaves inoculated with pEAQ-HT-GFP vector and infiltrated and dipped in the substrates mixture. **C**. Leaves inoculated with pEAQ-HT-GFP vector. **D**. Leaves inoculated with pEAQ-HT-DEST1:GAT1 construct and infiltrated and dipped in the substrates mixture. **E**. Leaves inoculated with pEAQ-HT-DEST1:GAT1 construct.

In the second experiment (Table 18), PAs and hydroxycinnamoyltartaric esters contents were not significantly different between control leaves and leaves transformed with *VvGAT1_Prt*. However, we observed a significant increase of %G (7.61 versus 6.31 in the control leaves). Free flavan-3-ol content was not impacted.

VvGAT1_Mac significantly increased hydroxycinnamoyltartaric esters (+20%) but decreased total PA (-33%) and ECG extension units in PAs compared to control leaves. %G was not significantly different in this condition. The leaves transiently transformed with *VvGAT2_Mac* tended to accumulate more PAs ($0.01 < P < 0.05$) and exhibited significantly more ECG as extension unit ($P = 0.006$). However, %G was not modulated. Moreover, the content of hydroxycinnamoyltartaric esters was significantly increased (+50%).

Table 17. Phenolic compounds of grapevine leaves inoculated with *Agrobacterium tumefaciens* (first experiment).

	GFP (control)	VvGAT1_Mac
Polymerised flavan-3-ols		
% G	1.45	4.74
Total content (mg.g^{-1} FW)	5.52	0.93
Hydroxycinnamoyl tartrates		
Total content (mg.g^{-1} FW)	1.5	0.79

Abbreviations are: FW: Fresh weight; %G: galloylation

Table 18. Phenolic compounds of grapevine leaves inoculated with *Agrobacterium tumefaciens* (second experiment).

	GFP (control)	VvGAT1_Prt	VvGAT1_Mac	VvGAT2_Mac
Polymerised flavan-3-ols				
ECG phloro	0.317 ± 0.007	0.343 ± 0.030	0.213 ± 0.008 **	0.510 ± 0.031 **
% G	6.31 ± 0.15	7.61 ± 0.25 **	6.3 ± 0.22	6.07 ± 0.36
Total content (mg.g^{-1} FW)	3.41 ± 0.10	3.06 ± 0.20	2.28 ± 0.11 **	5.73 ± 0.54 *
Free flavan-3-ols				
Catechin	0.038 ± 0.0029	0.030 ± 0.001 *	0.035 ± 0.002	0.052 ± 0.003 **
Epicatechin	0.005 ± 0.0003	0.008 ± 0.0012 *	0.006 ± 0.0004	0.016 ± 0.001 **
Total content (mg.g^{-1} FW)	0.043 ± 0.0032	0.038 ± 0.002	0.041 ± 0.002	0.068 ± 0.004 **
Hydroxycinnamoyl tartrates				
cis-caftaric acid	0.173 ± 0.007	0.207 ± 0.021	0.200 ± 0.005 **	0.255 ± 0.004 **
trans-caftaric acid	5.450 ± 0.238	6.373 ± 0.789	6.632 ± 0.126 **	7.758 ± 0.124 **
cis-coutaric acid	0.090 ± 0.004	0.133 ± 0.007 **	0.102 ± 0.001	0.179 ± 0.004 **
trans-coutaric acid	0.270 ± 0.011	0.393 ± 0.033	0.292 ± 0.003	0.572 ± 0.008 **
trans-fertaric acid	0.061 ± 0.003	0.081 ± 0.005 **	0.069 ± 0.003	0.089 ± 0.001 **
GRP	0.077 ± 0.018	0.164 ± 0.108	0.089 ± 0.011	0.304 ± 0.018 **
Total content (mg.g^{-1} FW)	6.122 ± 0.250	7.352 ± 0.748	7.384 ± 0.147 **	9.157 ± 0.133 **

Values presented are the mean of 3 independent extractions of the same sample ± standard deviation. FW: fresh weight.
Asterisks indicate that the mean values are statistically different (Student's t test) from those in the GFP transformed leaves (control).
*: Significance at $P < 0.05$; **: Significance at $P < 0.01$.
Abbreviations are: ECG phloro: epicatechin 3-*O*-gallate with phloroglucinol adduct; %G: galloylation rate; GRP: Grape reaction product (2-*S*-glutathionyl caftaric acid).

5. Discussion

We focused on 2 genes which encode putative SCPL-ATs (VvGAT1 and -2), identified as candidate genes for PA galloylation in grapevine by cross-data mining (Carrier *et al*., 2013).
Our phylogenetic study revealed that among the 51 SCPs encoded by grapevine genome, 12 genes putatively encode a SCPL-AT, including *VvGAT1* and *-2*. In grapevine genome, genes coding for VvSCPs are often physically grouped side by side or located near each other in genomic clusters. Their length ranges between 5 and 400 Kbp, depending on the number of genes and the non coding genomic regions that compose them. A reannotation of some genomic sequences was necessary. This study allows to discrimate between putative acyltransferase and peptidase functions for candidate genes annotated as serine carboxypeptidase in genomic data.

The main difference observed between real peptidases and SCPL-ATs was the replacement of Glu by Asp in the catalytic center. Stehle and co-workers (2006) observed this substitution in characterized Arabidopsis sinapoyltransferases. This substitution was observed for putative grapevine SCPL-ATs grouped in clade Ia. This residue could be involved in acylation reactions from glucose esters through glucose ester recognition.
The main phenolic esters found in grapevine are galloylated flavan-3-ols, hydroxycinnamoyltartrates and acylated anthocyanins. Due to its protein sequence identity (58%) with a characterized anthocyanin coumaroyltransferase (Noda *et al*., 2007), VvSCP31 is a good candidate for such activity in grapevine. However, our predictions are only based on phylogenetic analysis and we lack further clues concerning glucose ester and acyl acceptor affinity of these putative SCPL-ATs. The determination of the expression pattern in grape berry should be useful to establish hypothesis concerning their role *in planta* since phenolic esters found in grapevine have specific spatio-temporal accumulation profile during fruit development.

The previous attempts to produce active VvGATs failed in microorganisms (Fida Khater PhD). This time, VvGATs production was tested in plant organisms (tobacco and grapevine).

VvGAT coding sequence was cloned in the binary vector pEAQ-HT-DEST1 (Sainsburry *et al*., 2009) and incorporated in GV3101 strain of *A.tumefaciens*. Considering previous studies, no tag was added during VvGATs cloning (Mugford *et al*., 2010). Previous transient expression of SCPL-ATs in tobacco leaves allowed the characterization of the studied enzymes (Stehle *et al*., 2008, Mugford *et al*., 2009). In those studies, tobacco leaves were harvested at 5 and 6 days post inoculation (Stehle *et al*., 2008, Mugford *et al*., 2009, respectively).

SCPL-AT enzymatic activity was observed in MES buffer pH 6 (Stehle *et al*., 2008) or 6.5 (Mugford *et al*., 2009), MOPS buffer pH 7 (Shirley and Chapple, 2003), potassium phosphate buffer pH 6 (Lehfeldt *et al*., 2000). Maximal activity of SlGAC was measured at pH 5 in citrate-phosphate buffer (Li *et al*., 1999).
In spite of different assays, *VvGAT* transient expression in tobacco leaves did not yield proteins able to produce ECG from β-G and EC *in vitro*.
Different buffers (and pH) were used for galloylation assay: phosphate potassium buffer (pH 6), MES-KOH buffer (pH 5 and 6.5). The different pH conditions tested for galloylation *in vitro* are similar to those experimented in previous SCPL-AT functional validation studies.

Galloyltransferases involved in gallotannin biosynthesis had a pH optimum at 6 in citrate buffer (Fröhlich *et al*., 2002). From crude proteins extracts from tea leaves, epicatechin galloylation was optimal from pH 4 to 6 (Liu *et al*., 2012). Galloylation assay has been experimented from crude enzyme extract from pericarp of grapevine berries and reaction products will be analysed soon. This reaction could also be tested from crude enzyme extract from grapevine seeds or leaves using a larger range of pH and buffers. This should help to distinguish between the inactivity of enzymes produced in tobacco leaves or not optimized enzymatic reaction condition. In particular, if it is very likely that VvGATs use β-G as acyl donor, the chemical identity of acyl acceptors is still questionable. We can wonder if galloylation occurs on subunits of polymers after polymerization - which is doubtful due to tanning properties of large polymers on proteins- or on precursors before polymerization. The nature of the chemical species that are used by plant for elongation units is still a matter of debate (cf Chapter 1, Flavan-3-ol polymerization). In tea, the isolated enzyme is active on free monomers such as EC (Liu *et al*., 2012), but in this plant the galloylated flavan-3-ol pool is essentially composed of monomers and short polymers.

The relative expression level of *VvGAT* revealed that they are well expressed in the tobacco leaves inoculated with pEAQ-HT-DEST1:VvGAT constructs. Thus, the construct must be normally transferred in the host plant and expressed by its cellular machinery. The problem could come from the correct translation of *VvGAT* mRNA into protein *in planta*. The analysis of the total proteins extracted from tobacco leaves did not allow the identification of additional bands that could correspond to VvGATs. Nevertheless, similar results have been obtained with the same method for SCPL1 from oat (*Avena strigosa*) which was finally functional (Mugford *et al*., 2009). Tobacco leaves extracts should be probed with anti-VvGAT antisera to detect peptides corresponding to VvGAT. This more sensitive method allowed the detection of different bands corresponding to AsSCPL1 (Mugford *et al*., 2009). VvGAT-specific antisera would be useful for this experiment to check if the protein is produced and possibly cleaved.

SCPL-ATs can be subject to post-translational modifications such as glycosylation, internal cleavage of peptides and disulfide bridges. VvGAT1 and -2 carry 3 and 4 potential glycosylation sites, respectively. It has been demonstrated that the glycosylation by protein maturation mechanism may differ from one plant species to another (Gomord *et al*., 2010). VvGATs exhibit numerous potential *N*-glycosylation sites. It is possible that the tobacco cellular machinery does not have the ability to add the same oligosaccharide as that of the grapevine and produced a non-functional protein.
However, transient expression of SCPL-ATs: SMT from *Brassica napus* (Stehle *et al*., 2008) and SCPL1 from *Avena sativa* (Mugford *et al*., 2009) in tobacco leaves allowed to obtain functional proteins.

The impact of VvGAT overexpression on the phenolic compounds of tobacco leaves has been evaluated. A preliminary study was necessary to identify the different phenolic compounds found in tobacco leaves.
6 CQAs have been identified in tobacco leaves. Their biosynthesis can be ensured by HCT/HQTs as previously described in *Nicotiana tabacum* (Hoffmann *et al*., 2003, Niggeweg *et al*., 2004). Those enzymes have been identified as BAHD-ATs and are able to transfer caffeoyl moiety from caffeoyl-CoA to quinic acid. However, we have tested if CGAs could be synthesized from hydroxycinnamoyl glucose esters rather than hydroxycinnamoyl-CoA, as observed by Villegas & Kojima (1986) in *Ipomea batatas*. Hence, we have compared the content of the 4 main CQAs between non-inoculated tobacco leaves and leaves inoculated with pEAQ-HT-DEST1:VvGAT. VvGATs overexpression in

tobacco leaves led to concentration modifications of some hydroxycinnamoyl quinic acids. However, we observed a high variability of CQAs content among the plants transformed with the same genetic construct (3 biological replicates). This variability could be explained by differences in the efficiency of transient overexpression. Even if all the plants used for this experimented had the same age, we observed differences in growth rate and leaves development, that could explain such variability. Moreover, we will have to check if the suitable substrates of VvGAT (i.e. hydroxycinnamoyl glucose esters) are present in sufficient amount in tobacco leaves to catalyse such kind of reactions. These putative substrates have been purified from enzymatic reaction with VvGT2 (see Annexes, Supplemental Figure 7) and could be directly infiltrated in tobacco leaves to test in they influence CQA content in presence of VvGATs.

Tobacco leaves are rich in CQAs but devoid of phenolic compounds usually found in grapevine leaves, notably flavan-3-ols. As a result, the infiltration of substrates mixture (β-G and EC) was necessary to test galloylation reaction in tobacco leaves. ECG was not identified in tobacco leaves transformed with one of the VvGATs and infiltrated with the substrates. Substrates infiltration questions the fate of these molecules *in planta*. The infiltrated substrates (EC and β-G) were identified by UPLC-DAD-MS analysis. However, we do not know if those molecules are stored in the vacuole (or other intracellular compartments) where potentially occurs galloylation.

VvGAT1 was previously localized in pre-vacuolar vesicles in grapevine hairy-roots transformed with VvGAT1-GFP construct (Fida Khater PhD). Tobacco does not produce flavan-3-ols and probably does not possess the vesicular network necessary to import those metabolites into the vacuole. As a result, VvGAT is perhaps not addressed to this putative vesicular network *in planta*. It questions the fate of VvGATs fate in tobacco cells. Similarly, DgSCPLs described by Nishizaki and coworkers (2013) and suspected to be involved in anthocyanin acylation could not be functionally characterized using tobacco as heterologous organism to date (Nobuhiro Sasaki, personal communication).

Transient expression of VvGATs in grapevine leaves led to significant variations of phenolic compounds. Transformation with agrobacteria carrying *VvGAT1* cloned from Portan cultivar increased %G of PAs. Transient expression of VvGAT1 and -2 from Macabeu cultivar increased tartaric hydroxycinnamates content. Even if VvGATs have been tested as putative candidates for flavan-3-ol galloylation, their involvement in tartaric hydroxycinnamate biosynthesis is not excluded. Indeed, a previous study demonstrated that Arabidopsis SCPL17 could catalyse both benzoylation and sinapoylation of glucosinolates (Lee *et al*., 2012). We lack further information concerning the amino acids that could discriminate acyl donor and/or acceptor.
The relative expression level of *VvGAT* must be experimented to check if the introduced gene is really overexpressed in grapevine leaves. Other parameters could influence VvGATs activity *in planta*. Notably, the content of acyl donors (β-G and hydroxycinnamoyl glucose esters) in leaves must influence VvGATs activity. The transport of those putative substrates in the adequate compartment of the cell may also be a limiting factor. To cope with this, substrates could be mixed to agrobacteria solution before infiltration under vacuum.
The results obtained by grapevine leaves agroinfiltration are promising but should be considered as primary results and need to be repeated.

CHAPTER 5: GENERAL DISCUSSION AND PROSPECTS

A better understanding of the molecular mechanisms involved in biosynthesis, transport and storage of compounds of oenological interest in grape berry is necessary to control the quality of this fruit. Flavan-3-ols are the most abundant phenolic compounds in grape berry and influence bitterness, astringency and color stability of wine. Otherwise, pharmacological studies have shown that these molecules have beneficial properties for human health. The flavan-3-ols found in grapevine have the particularity to be partially acylated with GA. The acylation reaction involved is called galloylation and influences oenological and pharmacological properties of those molecules.
Flavan-3-ols are widely found in plant kingdom and their biosynthesis has been extensively studied. Most of the genes responsible for flavan-3-ols biosynthesis have been discovered in plants and notably Arabidopsis and grapevine. However, flavan-3-ol galloylation seems to be restricted to few plants, including grapevine. As a result, it is a suitable plant to study the genetic bases and the molecular mechanisms involved in the biosynthesis of those secondary metabolites. A gene involved in GA biosynthesis has been highlighted in walnut only. The molecular mechanisms able to influence galloylation rate of PAs *in planta* have been studied in a restricted number of species which accumulate galloylated flavan-3-ols (persimmon, tea plant). However, in spite of the identification of genes potentially involved in this mechanism, none of them has been functionally characterized.

This work fits into a recently emerging field: phytochemical genomics (Saito, 2013). This new trend tends to identify the genetic bases involved in the biosynthesis of plant metabolites (phytochemicals), combining different –omics strategies, to establish a gene to metabolite link. Indeed, cross-data mining is an efficient strategy to restrict the number of candidate genes before functional validation. In Arabidopsis, phytochemical genomics is accelerated by the existence of numerous mutants and database of –omics data of every-day growing size. In grapevine, the generation of mutants could be possible but would require a long-term process, as well as space and labor to cultivate them. For this work, we have combined previous genetic (QTL mapping and association genetics) and transcriptomic data with metabolomics (HPLC-MS) to establish a gene to metabolite link.

The studied metabolites were flavan-3-ols and GA, necessary to the biosynthesis of galloylated flavan-3-ols.
The aim of this work was to validate the function predicted to selected candidate genes.
Four shikimate dehydrogenases (VvSDHs) have been studied for their capacity to produce GA from shikimate pathway metabolites.
Two serine carboxypeptidases-like (VvGATs) have been studied for their capacity to acylate flavan-3-ols with galloyl moiety.
Our hypothesis predicted that several genes were necessary to encode enzymes involved in a multistep galloylation of flavan-3-ols. Indeed, we hypothesized that shikimate dehydrogenase(s) could produce GA, further glucosylated by glucosyltransferases (VvGTs) to form β-G. This last molecule may be a precursor for the transfer of galloyl moiety on flavan-3-ols catalyzed by VvGAT(s). Otherwise, due to the capacity of VvGTs to produce hydroxycinnamoyl glucose esters, the study of VvGATs has been extended to tartaric acid acylation with hydroxycinnamoyl moiety from their glucose ester.

The expression pattern of those genes in grape berry tissues, *in vitro* and *in planta* activities allowed us to accumulate several arguments in favor of a two-step galloylation of flavan-3-ols.

Strategies to validate candidate genes

The knowledge concerning grapevine metabolites and the analysis methods previously developed in the host laboratory allowed to bring information to answer the initial scientific question.

The heterologous expression of VvSDHs in *E.coli* quickly provided active enzymes that have been characterized *in vitro*. *In planta* validation resulted from a longer process which consisted in stable transformation of hairy-roots. The generation of transgenic hairy-roots is a strategy routinely employed in the laboratory to validate the function of candidate genes such as flavonoid TFs (Terrier *et al*., 2009, Cutanda-Perez *et al*., 2009, Huang *et al*., 2014) and anthocyanin transporters (Gomez *et al*., 2009, 2011).

The study of VvGATs required the use of plants rather than microorganisms. Indeed, earlier heterologous expression in *E.coli* and 2 yeast species did not lead to active enzymes. Plant leaves (tobacco and grapevine) transient transformation was adopted for this work. Transient expression of genes in grapevine leaves inoculated with *Agrobacterium tumefaciens* was previously developed in INRA Colmar.

The development of a greenhouse in the host laboratory allowed me to grow plant material (tobacco, grapevine plantlets) necessary for transient transformation. Tobacco growth rate allows getting workable leaves 6 weeks after sowing. On the contrary, grapevine vitroplant growth is relatively slow and workable about 3 months after the first cuttings. As a result, vitroplant availability has been a limiting factor for GAT study.
The culture and grounding of plant material (tobacco, hairy-roots) was more time consuming but necessary to provide strong arguments for gene validation.

Otherwise, I have benefited from previous works on grape berries harvested in Montpellier SupAgro campus vineyard.

The Polyphenol Platform helped me to determine the metabolite profile of plant tissues with high precision. Notably, I could acquire precious data concerning GA and β-G content in grape berry tissues and transgenic hairy-roots. Classic HPLC methods were previously developed to assay molecules highly concentrated in grape berries (flavonoids, hydroxycinnamoyltartrates) but were not sensitive enough to detect trace amounts of GA and β-G. UHPLC-QqQ-MS method allowed to quantify low concentrations of these metabolites, down to 1 $nmol.g^{-1}$ for GA in grape berry pulp. For the first time in grapevine, β-G content has been quantified in the main tissues at key developmental stages of the grape berry.

Genetic bases of GA biosynthesis and galloylation

The detection of QTLs which modulate PA %G attests the existence of genes responsible for galloylation of flavan-3-ols. Candidate genes are localized under QTL intervals which influence PA %G in grape berry skin. However, this trait is probably controlled by many genomic portions and digenic epistatic interactions (Huang *et al*., 2012) and genetic study using other populations would probably have revealed other genomic regions. The discovery of numerous PA specific TFs, which

regulate positively or negatively PAs biosynthesis in grapevine, attests for a complex regulatory mechanism. Our candidate genes are induced by several positive TFs (MybPA1 and/or MybPA2, and MybPAR).

The tested candidate genes are often grouped side by side (*VvSDH7, -8* and *-9*; *VvGT1*, *-2* and *-3*) or closely (*VvGAT1* and *-2*) on the same chromosome. This proximity must result from gene duplication. In plants, homologous genes are often found in gene clusters resulting from tandem duplication. Gene duplicates are likely responsible for the diversification of phytochemicals. About 15-20% of genic content is composed of tandem gene clusters in rice and Arabidopsis (Rizzon *et al*., 2006). Several paralogs from a common ancestral gene can be involved in the biosynthesis of the same metabolite, with more or less efficiency. Briefly, gene duplication can result in loss of function (pseudogenisation), subfunctionalization, or in the acquisition of a new function (neofunctionalization) (Barker *et al*., 2012). The case of terpene synthase family in Norway spruce is a good illustration of neofunctionalization (Keeling *et al*., 2008). Few key amino acids substitutions have allowed the diversification of substrate affinity and terpenoid production.
If the ancestral gene assumed several functions, they can be partitioned among the new paralogs (protein subfunctionalization). A duplicated gene can also acquire an organ-specific expression pattern (expression subfunctionalization). Interestingly, duplication can result in protein subcellular localization, associated with signal peptide (reviewed by Byun-McKay & Geeta, 2007). For example, Ding and coworkers (2007) propounded that the loss of the plastidic transit peptide in NtSDH2 sequence could delocalize the shikimate pathway in the cytosol or result in a new function.

3 *VvSDHs* were localized in a 34.1 kb cluster in chromosome 14 whereas *VvSDH5* is isolated on chromosome 5. A QTL which influences PAs %G in berry skin has been detected on chromosome 14.
Tandem gene duplication has been highlighted for SDH genes in woody species, including grapevine (Tohge *et al*., 2013). Comparing key amino acid motifs in the SDH domain of the unique SDH of Arabidopsis with Dicots SDHs, we observed common divergences. Those substitutions could affect substrate binding and orientation, and thus GA biosynthesis.

We identified 12 putative SCPL-ATs, including VvGAT1 and -2, in the grapevine genome. Two independent clusters composed of 7 and 3 putative SCPL-ATs were localized on chromosome 3 and 11, respectively. Two single genes encoding a putative SCPL-AT were also localized on chromosome 11 and 18.
A single mutation in the catalytic center could explain that some SCPs acquired the capacity to transfer acyl groups from glucose esters, and lost their ancestral peptidase activity.
VvGAT1 and *-2* co-localize in chromosome 3 with 5 other putative SCPL-ATs in a ~400 kb cluster.
A large QTL interval which influences PA %GA in berry skin has been detected in this region on chromosome 3.

The putative SCPL-ATs must result of tandem gene duplication according to the proximity of the genes on the same chromosome and their high level of sequence identity. Three VvGTs able to produce β-G are also grouped on chromosome 3, in the vicinity of VvGATs and putative SCPL-ATs. The co-localization of VvGT, VvGATs and putative SCPL-ATs on chromosome 3 lets think about the existence of a functional cluster involved in secondary metabolites biosynthesis. However, VvGTs and VvGATs are separated by ~1611 kb. Some gene clusters involved in plant secondary metabolism in plants were already identified. But their size is smaller since they are constituted of 3-10 genes found in ~35–270 kb (Nützmann & Osbourn, 2014).

VvGTs can provide a large range of HBA and HCA glucose-esters (Khater *et al*., 2012). Putative SCPL-ATs could use those glucose esters as substrates to catalyze acylation reactions. However, the specificity of these putative SCPL-ATs cannot be determined from *in silico* analysis and requires functional validation. Indeed, few SCPL-ATs have been characterized and the existence of key motifs which discriminate acyl donor or acceptor molecules have not been highlighted yet. As a result, it is not excluded that these putative SCPL-ATs would be able to use different kind of glucose ester for their acyltransferase activity.

Functional validation of candidate genes

The capacity of a walnut SDH to produce GA from SA pathway metabolites (Muir *et al*., 2011) led us to investigate the function of the 4 SDHs encoded in grapevine genome. Indeed, since the publication of this study, GA biosynthesis has no more been investigated in plants.
Escherichia coli was a suitable organism to produce VvSDHs. Indeed, we rapidly obtained active enzymes. Among the 4 VvSDHs, VvSDH5 exhibits the higher "classical" SDH activity from SA/$NADP^+$ and 3-DHS/NADPH. VvSDH8 and -9 produce the higher GA amount from 3-DHS and $NADP^+$. VvSDH7 exhibits very low "classical" SDH activity and produced few GA from 3-DHS.

VvSDH8 overexpression in grapevine hairy-roots enhanced GA metabolism and confirms its involvement in GA biosynthesis *in planta*. In transgenic hairy-roots, galloylated flavan-3-ols content was significantly higher compared to control line. GA was likely glucosylated to increase β-G pool and allow the transfer of galloyl moieties to flavan-3-ols. Genetic manipulation triggered a leak in the shikimate pathway carbon flux and enhanced GA metabolism.

The biosynthesis of galloylated flavan-3-ols has been mainly studied in persimmon and tea plant to improve fruit and leaves quality, respectively. Persimmon PAs are responsible for non desirable astringency of fruits. Tea plant leaves are rich in galloylated catechins which exhibit health beneficial properties and subject to numerous pharmacological studies. In tea plant, they are highly accumulated in buds and young leaves, but absent in roots (Jiang *et al*., 2013). The biosynthesis of galloylated flavan-3-ols was only shown from crude extracts of leaves (Liu *et al*., 2012). On the contrary, the biosynthesis of hydroxycinnamoyltartrates was very little investigated (Sullivan, 2014). Nevertheless, several studies report the biosynthesis of structurally close esters composed of a hydroxycinnamoyl moiety esterified by an organic acid. Sinapoylmalate in Arabidopsis (Lehfeldt *et al*., 2000) and caffeoylmalate in red clover (Sullivan, 2009) can be synthesized by SCPL and BAHD-AT, respectively.

Tobacco leaves transient transformation did not allow production of active enzymes when tested *in vitro*. The galloylation reaction did not occur *in planta* when putative substrates (β-G and EC) were co-infiltrated.
However, the infiltration of agrobacteria carrying VvGATs in grapevine leaves triggered variations of galloyated flavan-3-ols and/or hydroxycinnamoyl tartrates.

As SCPL-ATs can be subject to glycosylation and/or internal cleavage, it is possible that tobacco cells do not possess the suitable mechanisms necessary to produce an active acyltransferase. Indeed,

VvGAT1 and -2 exhibit 3 and 4 *N*-glycosylation sites, respectively. Glycosylation tag can differ between plants and could be different between tobacco and grapevine (Gomord *et al*., 2010).

Transient expression in homologous rather than in heterologous organism is probably more suitable for VvGATs characterization.
Transient expression of Portan VvGAT1 in grapevine leaves triggered a significant increase of PAs %G and tended to increase the content of tartaric hydroxycinnamoyltartrates.
Macabeu VvGAT1 and -2 significantly increased tartaric hydroxycinnamoyltartrates content.
It is not excluded that VvGATs can be responsible for both flavan-3-ols galloylation and transfer of hydroxycinnamoyl moieties on tartaric acid. Indeed, SCPL17 was proposed to catalyze both benzoylation and sinapoylation of glucosinolates in Arabidopsis (Lee *et al*., 2012).

Correlations between metabolites profile and expression pattern of candidate genes in grape berry.

The expression level of the studied genes in grape berry is a good strategy to determine if the encoded enzymes are synchronously produced with the studied phenolic compounds, and in the same tissues.
This strategy has been frequently employed to study other genes of the flavonoid pathway (e.g. Bogs *et al*., 2005, Terrier *et al*., 2009).
As galloylated flavan-3-ols are synthesized during the green stage, the expression of genes involved in this mechanism is expected during phase I.
GA content in grape berry tissues was seldom investigated (Yilmaz & Toledo, 2004). The results suggest that the pool size of GA is limited compared to the production of flavan-3-ols in grape berry.
GA and β-G are also biosynthesized in green stage, mainly in seeds and in skin, to a lesser extent.
All *VvSDHs* followed a similar expression pattern in pericarp during the green stage with a peak of expression at 35 daf. *VvSDH5* exhibited the higher relative expression whatever the developmental stage and the tissue. Its expression stayed relatively high few days after *véraison*. This enzyme must ensure the carbon flux in the shikimate pathway, necessary for the biosynthesis of primary metabolites such as phenylalanine.
On the contrary, *VvSDH7* was hardly expressed in grape berry. If its predicted quinate dehydrogenase activity is confirmed, its low expression could explain the absence of quinic acid and derivatives in *Vitis vinifera*.
VvSDH8 and *-9* expressions are synchronous with GA accumulation. Contrary to VvSDH9, VvSDH8 must be a seed-specific enzyme, reflecting an expression subfunctionalization of this gene.

Once synthesized, GA must be activated by glucosylation to an energy rich molecule, necessary for the transfer of the galloyl moiety on flavan-3-ols. This step represents a second limiting factor to produce galloylated flavan-3-ols. The content of β-G, the putative precursor for flavan-3-ol galloylation, was investigated for the first time in grape berry. The highest content of β-G has been detected in immature seeds. It could result of the activity of previously characterized VvGTs involved

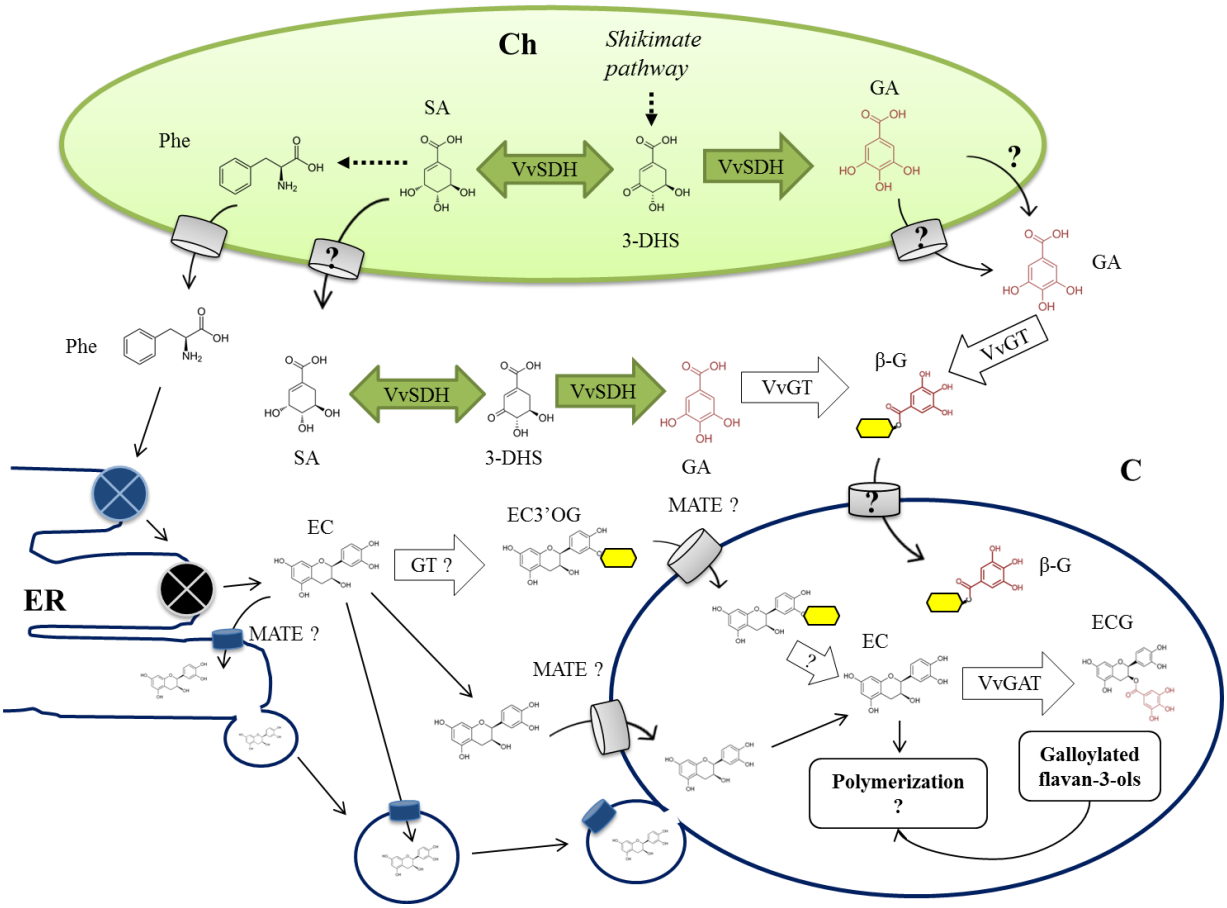


Figure 42. A model for flavan-3-ol galloylation in grapevine.
Transporters are represented by cylinders. Metabolons are represented by crossed circles with in blue phenylpropanoid pathway and in black flavonoid pathway. Glucosyl moiety is respreseted by yellow hexagons. Gallic acid and galloyl moiety are coloured in red.
Abbreviations: **Ch**: chloroplast, **C**: cytosol, **ER**: endoplasmic reticulum, **N**: nucleus, SA: shikimic acid, 3-DHS: 3-dehydroshikimate, GA: gallic acid, β-G: β-glucogallin, EC: epicatechin, EC3'OG: epicatechin 3'-*O*-glucoside, Phe: phenylalanine, VvSDH: *Vitis vinifera* Shikimate dehydrogenase, VvGT: *Vitis vinifera* Glucosyltransferase, VvGAT: *Vitis vinifera* Glucose Acyltransferase, **?**: unknown actor.

in GA glucosylation (Khater *et al*., 2012) and could explain why PAs are more galloylated in seeds than skin and pulp. Like *VvGATs*, *VvGTs* were mainly expressed during the green stage.
In pericarp, *VvGATs* exhibited a peak of expression at 11 daf, and their expression decreased to *véraison*. *VvGAT1* was mainly expressed in seeds and skin during the green stage, that fitted with galloylated PAs accumulation. Its relative expression was also high in mature skin (99 daf), suggesting its possible involvement in anthocyanins acylation also. *VvGAT2* was mainly expressed in the 3 tissues during the green stage.

Globally, the expression pattern of these genes is synchronous with PA accumulation in grape berry, suggesting that all these enzymes could be involved in galloylation. When expressed in immature pulp and mature skin, VvGATs could be involved in acylation of tartaric acid and anthocyanins, respectively. Otherwise, one can wonder why seeds contain more galloylated PAs than the other grape berry tissues. Those genes could be regulated by a specific regulation system, in addition to those regulating the PA biosynthetic pathway.

A model for flavan-3-ol galloylation in grapevine

Different molecular mechanisms potentially involved in flavan-3-ol galloylation have been represented at the cellular level (Fig. 42). Combining our results with previous studies concerning the biosynthesis, transport and storage of different molecules involved in this mechanism, different putative pathways have been modeled.
Considering that the prediction of putative plastidic signal peptide is not always reliable, the subcellular localization of VvSDHs (chloroplast or cytosol) remains to be determined.
However, SA pathway being recognized as plastidic, a chloroplastic VvSDH could benefit of the abundance of SA pathway metabolites to bypass this pathway and produce GA.
A cytosolic biosynthesis of GA is not excluded and has been proposed (Ossipov *et al*., 2003). It could favor the spatial colocalization of GA with VvGTs to produce more rapidly β-G, the substrate of VvGATs. However, this hypothesis would require the export of shikimate pathway intermediates (e.g. SA) across the chloroplast envelop membranes. The transfer of shikimate pathway metabolites (SA, 3-DHS) through the chloroplast envelop was demonstrated in isolated chloroplasts (Leuschner & Schultz, 1991).
Once synthesized in the cytosol, β-G must be routed towards the compartment where galloylation occurs. The transporter(s) involved in the transport of β-G remain to be identified.

The shikimate pathway leads to the production of phenylalanine that can be exported towards the cytosolic compartment. Phenylalanine plastidic transporter has recently been highlighted (Widhalm *et al*., 2015). Phenylalanine is used by the enzymes of the phenylpropanoid/flavonoid pathway grouped in metabolons at the cytosolic face of the endoplasmic reticulum.
Several models have been proposed for flavan-3-ols routing towards vacuolar compartment (reviewed by Zhao, 2015). Three main mechanisms, mediated by transporters, vesicles or GST, could collaboratively take part in flavan-3-ols subcellular distribution. In *Medicago truncatula*, the glucosylation of epicatechin by a specific glucosyltransferase yields epicatechin 3'-*O*-glucose, then transported by a MATE-type transporter to the vacuole (Pang *et al*., 2008, 2013). In grapevine, epicatechin glucoside has been detected in seeds (Delcambre & Saucier, 2012). However, the molecular actors involved in glucosylation and transport have not been described in grapevine. Two grapevine MATE-type transporters, VvMATE1 and -2, have been localized at the tonoplast and the

Golgi apparatus, respectively. Their expression during green stage let think that they could be involved in flavan-3-ols transport but their substrates specificity (free or glucosylated flavan-3-ols) has not been determined.

VvGATs are predicted to be addressed to intracellular compartment. In hairy-roots transformed with VvGAT1:GFP construct, VvGAT1 was localized in vesicles, that could be pre-vacuolar vesicles (Fida Khater PhD). These vesicles could be involved in the transport of PAs towards the vacuole (Zhao *et al*., 2010). The presence of β-G and flavan-3-ols (or precursors) in the same compartment is required for VvGAT(s) to catalyze galloylation. However, the exact compartment where flavan-3-ol galloylation occurs is still unclear.

Galloylation probably occurs on free flavan-3-ols before PAs polymerization. Indeed, long chains of PAs have the capacity to interact with proteins and precipitate them and thus inhibit their activity. The low amount of catechin 3-*O*-gallate compared to the abundance of ECG in grapevine lets think that epicatechin (or one of its precursor) is likely the monomer subject to galloylation.

Short-term prospects

It would be interesting to model the interaction of key amino acids with the substrate for each group of DQD/SDH. We suppose that the observed divergences compared to Arabidopsis SDH must influence catalytic properties. Site-directed mutagenesis targeted on those amino acids followed by *in vitro* study could be used to check our hypotheses.

The impact of the 3 other *VvSDHs* overexpression in grapevine hairy-roots must be investigated to confirm their predicted role from *in vitro* analysis. In spite of the different trials, we generated only one hairy-root line transformed with *VvSDH9*. The phenolic composition of this line will be analyzed to confirm the role of this gene *in planta*. The high GA production observed for this enzyme *in vitro* could result in a high GA content potentially toxic for the cells.
The intracellular localization of VvSDHs could be investigated in the generated grapevine hairy-roots transformed with VvSDH:GFP construct.
The same strategy could be used to localize VvGTs, expected to be cytosolic enzymes, and should be pursued for VvGATs.

The relative expression levels of the 10 other putative grapevine SCPL-ATs could be investigated along grape berry development and in the main tissues. Hence, we could hypothesize if they are involved in the biosynthesis of phenolic compounds esters in green stage (galloylated flavan-3-ols, hydroxycinnamoyltartrates) or later during ripening (acylated anthocyanins).

Long-term prospects

In planta validation of the studied genes could be monitored by the stable transformation of a whole grapevine vitroplant.

The hypothesis of a two-step galloylation of flavan-3-ols must involve several genes encoding transporters which allow the precursors (GA, β-G, flavan-3-ols) to cross cellular membranes. These transporters remain to be identified.

VvSDHs and *VvGATs* genes could be used to develop genetic markers relative to %G of PAs in future cultivar creation, better adapted to viticulture constraints such as pathogens resistance and berry quality. Indeed, the abundance of galloylated flavan-3-ols at maturity must be controlled due to the influence of %G on wine astringency. Otherwise, galloylated flavan-3-ols could be involved in defense mechanisms against pathogens in leaves as demonstrated in tea plant. Controlling the %G of leaf PA independently of berry composition could be a challenge for sustainable viticulture. The identification of DQD/SDH orthologs in dicots opens the way for the biochemical characterization of SDHs involved in GA biosynthesis. Notably it would be interesting to transform monocotyledonous plants or GA-poor plants (e.g. Arabidopsis) with a VvSDH able to produce GA.

The mechanisms investigated for flavan-3-ols galloylation in grapevine could be transposed to plants which produce those metabolites (persimmon, tea plant), hydrolysable tannins or other GA derivatives.
The hypothesis of a two-step galloylation of flavan-3-ols must involve other molecular actors (e.g. TFs, transporters). They remain to be identified and characterized to complete the puzzle. Genome-wide association (GWA) study using a core-collection of grapevine cultivars has been started in collaboration with other researchers. It must be a good strategy to highlight other candidate genes. This study could also be useful to identify genetic polymorphisms among the 12 putative grapevine SCPL-ATs and determine if any genetic variant can be associated to particular phenolic ester accumulation.

References

Food and Agriculture Organization (FAO): http://faostat3.fao.org/home/E.

International Organization of Vine and Wine (OIV): http://www.oiv.int/oiv/cms/index

R Development Core Team. 2009. R: a language and environment for statistical computing, Vienna, Austria: R Foundation for Statistical Computing.

Abbasi AM, Guo X, Fu X, Zhou L, Chen Y, Zhu Y, Yan H, Liu RH. 2015. Comparative assessment of phenolic content and in vitro antioxidant capacity in the pulp and peel of mango cultivars. International Journal of Molecular Sciences **16**:13507–27.

Abrahams S, Tanner GJ, PJ Larkin, AR Ashton. 2002. Identification and biochemical characterization of mutants in the proanthocyanidin pathway in Arabidopsis. Plant Physiology **130**: 561–576.

Abrahams S, Lee E, Walker AR, Tanner GJ, Larkin PJ, Ashton AR. 2003. The Arabidopsis TDS4 gene encodes leucoanthocyanidin dioxygenase (LDOX) and is essential for proanthocyanidin synthesis and vacuole development. The Plant Journal **35**: 624–636.

Achnine L, Huhman DV, Farag MA, Sumner LW, Blount JW, Dixon RA. 2005. Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*.The Plant Journal **41**: 875–887.

Adams DO. 2006. Phenolics and ripening in grape berries. American Journal of Enology and Viticulture **57**: 249–256.

Adrian M, Jeandet P, Duillet-Breuil AC, Tesson L, Bessis R. 1997. Biological activity of resveratrol, a stilbenic compound from grapevines, against *Botrytis cinerea*, the causal agent for gray mold. Journal of Chemical Ecology **23**: 1689–1702.

Agasse A, Vignault C, Kappel C, Conde C, Geros H, Delrot S. 2009. Sugar transport and sugar sensing in grape. In: Roubelakis-Angelakis KA, ed. Grapevine molecular physiology and biotechnology. New York: Springer, 105–128.

Agati G, Azzarello E, Pollastri S, Tattini M. 2012. Flavonoids as antioxidants in plants: location and functional significance. Plant Science **196**: 67–76.

Ageorges A, Fernandez L, Vialet S, Merdinoglu D, Terrier N, Romieu C. 2006. Four specific isogenes of the anthocyanin metabolic pathway are systematically co-expressed with the red colour of grape berries. Plant Science **170**: 372–383.

Akagi T, Ikegami A, Tsujimoto T, Kobayashi S, Sato A, Kono A, Yonemori K. 2009. DkMyb4 is a Myb transcription factor involved in proanthocyanidin biosynthesis in persimmon fruit. Plant Physiology **151**: 2028–2045.

Akagi T, Ikegami A, Yonemori K. 2010. DkMyb2 wound-induced transcription factor of persimmon (*Diospyros kaki* Thunb.), contributes to proanthocyanidin regulation. Planta **232**: 1045–1059.

Alfenito MR, Souer E, Goodman CD, Buell R, Mol J, Koes R, Walbot V. 1998. Functional complementation of anthocyanin sequestration in the vacuole by widely divergent glutathione S-transferases. Plant Cell **10**: 1135–1149.

Alonso-Amelot ME, Oliveros-Bastidas A, Calcagno-Pisarelli MP. 2007. Phenolics and condensed tannins of high altitude *Pteridium arachnoideum* in relation to sunlight exposure, elevation, and rain regime. Biochemical Systematics and Ecology **35**: 1–10.

Amrhein N, Deus B, Gehrke P, Steinrücker H. 1980. The site of the inhibition of the shikimate pathway by glyphosate. II. Interference of glyphosate with chorismate formation *in vivo* and *in vitro*. Plant Physiology **66**: 830–834.

An C, Mou Z. 2011. Salicylic acid and its function in plant immunity. Journal of Integrated Plant Biology **53**: 412–28.

Andersen OM, Jordheim M. 2006. The Anthocyanins. In: Andersen, O.M. and Markham, K.R., Eds., Flavonoids Chemistry, Biochemistry and Applications, CRC Press, Taylor and Francis, Boca Raton, 471–551.

Anke J, Petereit F, Engelhardt C, Hensel A. 2008. Procyanidins from *Myrothamnus flabellifolia* Welw. Natural Product Research **22**: 1243–1254.

Appelhagen I, Nordholt N, Seidel T, Spelt K, Koes R, Quattrochio F, Sagasser M, Weisshaar B. 2015. TRANSPARENT TESTA 13 is a tonoplast P3A -ATPase required for vacuolar deposition of proanthocyanidins in *Arabidopsis thaliana* seeds. Plant Journal **82**: 840–9.

Arnold RA, Noble AC, Singleton VL. 1980. Bitterness and astringency of phenolic fractions in wine**.** Journal of Agricultural and Food Chemistry **28**: 675-678.

Ayaz FA, Hayırlıoglu-Ayaz S, Alpay-Karaoglu S, Grúz J, Valentová K, Ulrichová J, Strnad M. 2008. Phenolic acid contents of kale (*Brassica oleraceae* L. var. *acephala* DC.) extracts and their antioxidant and antibacterial activities. Food Chemistry **107**: 19–25.

Barceló AR, Pomar F, López-Serrano M, Pedreño MA. 2003. Peroxidase: a multifunctional enzyme in grapevines. Functional Plant Biology 30: 577–591.

Barker MS, Baute GJ, Liu SL. 2012. Duplications and Turnover in Plant Genomes. In: Wendel *et al*. (eds), Plant Genome Diversity Volume 1, Springer-Verlag Wien, 155-169.

Bate-Smith EC. 1968. The phenolic constituents of plants and their taxonomic significance. Journal of the Linnean Society of London, Botany **60**: 325–356.

Bate-Smith EC. 1972. Chemistry and phylogeny of the angiosperms. Nature **269**: 353-354.

Bate-Smith EC. 1973. Systematic distribution of ellagitannins in relation to the phylogeny and classification of the angiosperms. In: Bendz G, Santesson J (eds) Chemistry in botanical classification. Academic Press London, 93-102.

Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L. 2004. TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. Plant **39**: 366–380.

Baur JA, Sinclair DA. 2006. Therapeutic potential of resveratrol: the *in vivo* evidence. Nature Reviews Drug Discovery **5**: 493–506.

Bavaresco L, Fregoni M, Trevisan M, Mattivi F, Vrhovsek U, Falchetti R. 2002. The occurrence of the stilbene piceatannol in grapes. Vitis **41**:133–136.

Bavaresco L, Pezzutto S, Gatti M, Mattivi F. 2007. Role of the variety and some environmental factors on grape stilbenes. Vitis **46**: 57–61.

Baxter IR, Young JC, Armstrong G, Foster N, Bogenschutz N, Cordova T, Peer WA, Hazen SP, Murphy AS, Harper JF. 2005. A plasma membrane H^+-ATPase is required for the formation of proanthocyanidins in the seed coat endothelium of *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America **102**: 2649–2654.

Becerril J, Duke S, Lydon J. 1989. Glyphosate effects on shikimate pathway products in leaves and flowers of velvetleaf. Phytochemistry **28**: 695–700.

Bell-Lelong DA, Cusumano JC, Meyer K, Chapple C. 1997. Cinnamate-4-hydroxylase expression in Arabidopsis. regulation in response to development and the environment. Plant Physiology **113**: 729–38.

Bertolini A, Peresson C, Petrussa E, Braidot E, Passamonti S, Macri F, Vianello A. 2009. Identification and localization of the bilitranslocase homologue in white grape berries (*Vitis vinifera* L.) during ripening. Journal of Experimental Botany **60**: 3861–71.

Bianco A, Buiarelli F, Cartoni G, Coccioli F, Jasionowska R, Margherita P. 2003. Analysis by liquid chromatography-tandem mass spectrometry of biophenolic compounds in olives and vegetation waters, part 1. Journal of Separation Science **26**: 409–416.

Bicker J, Petereit F, Hensel A. 2009. Proanthocyanidins and a phloroglucinol derivative from *Rumex acetosa* L. Fitoterapia **80**: 483–495.

Bischoff M, Schaller A, Bieri F, Kessler F, Amrhein N, Schmid J. 2001. Molecular characterization of tomato 3-dehydroquinate dehydratase-shikimate:NADP oxidoreductase. Plant Physiology **125**: 1891–1900.

Blouin J, Cruège J. 2003. Analyse et Composition des Vins: Comprendre le Vin, Editions La Vigne, Dunod, Paris, France, 304 pp.

Boerjan W, Ralph J, Baucher M. 2003. Lignin biosynthesis. Annual Review of Plant Biology **54**: 519–46.

Bogs J, Downey MO, Harvey JS, Ashton AR, Tanner GJ, Robinson SP. 2005. Proanthocyanidin synthesis and expression of genes encoding leucoanthocyanidin reductase and anthocyanidin reductase in developing grape berries and grapevine leaves. Plant Physiology **139**: 652–663.

Bogs J, Ebadi A, McDavid D, Robinson SP. 2006. Identification of the flavonoid hydroxylases from grapevine and their regulation during fruit development. Plant Physiology **140**: 279–291.

Bogs J, Jaffé FW, Takos AM, Walker AR, Robinson SP. 2007. The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. Plant Physiology **143**: 1347–1361.

Bonner CA, Jensen RA. 1994. Cloning of cDNA encoding the bifunctional dehydroquinase shikimate dehydrogenase of aromatic amino-acid biosynthesis in *Nicotiana tabacum*. Biochemical Journal **302**: 11–14.

Bontpart T, Cheynier V, Ageorges A, Terrier N. 2015. BAHD or SCPL acyltransferase? What a dilemma for acylation in the world of plant phenolic compounds. New Phytologist **208**: 695–707.

Borkowski T, Szymusiak H, Gliszczyńska-Swigło A, Tyrakowska B. 2005. The effect of 3-*O*-β-glucosylation on structural transformations of anthocyanins. Food Research International **38**: 1031–1037.

Boss PK, Davies C, Robinson SP. 1996. Expression of anthocyanin biosynthesis pathway genes in red and white grapes. Plant Molecular Biology **32**: 565-9.

Boss PK, Davies C. 2001. Molecular biology of sugar and anthocyanin accumulation in grape berries. In: Roubelakis-Angelakis KA (Eds) Molecular Biology and Biotechnology of the Grapevine, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 1–33.

Boss PK, Davies C. 2009. Molecular biology of anthocyanin accumulation in grape berries. In: Roubelakis-Angelakis KA, ed. Grapevine molecular physiology and biotechnology. New York: Springer, 263–287.

Böttcher C, Keyzers R, Boss P, Davies C. 2010. Sequestration of auxin by the indole-3-acetic acid-amido synthetase GH3–1 in grape berry (*Vitisvinifera* L.) and the proposed role of auxin conjugation during ripening. Journal of Experimental Botany **61**: 3615–3625.

Boukharta M, Girardi M, Metche M. 1988. Procyanidines galloylées du sarment de vigne (*Vitis vinifera*). Séparation et identification par chromatographie liquide haute performance et chromatographie en phase gazeuse. Journal of Chromatography **455**: 406–409.

Boulekbache-Makhlouf L, Meudec E, Chibane M, Mazauric JP, Slimani S, Henry M, Cheynier V, Madani K. 2010. Analysis by high-performance liquid chromatography diode array detection mass spectrometry of phenolic compounds in fruit of *Eucalyptus globulus* cultivated in Algeria. Journal of Agricultural and Food Chemistry **58**:12615-24.

Boursiquot JM, Sapis JC, Macheix JJ. 1986. Les esters hydroxycinnamiques chez le genre Vitis. Essai d'application taxonomique: premiers résultats. Comptes Rendus Académie des Sciences **302**: 177.

Braidot E, Petrussa E, Bertolini A, Peresson C, Ermacora P, Loi N, Terdoslavich M, Passamonti S, Macri F, Vianello A. 2008. Evidence for a putative flavonoid translocator similar to mammalian bilitranslocase in grape berries (*Vitis vinifera* L.) during ripening. Planta **228**: 203–213.

Breddam K, Sørensen SB. 1987. Isolation of carboxypeptidase III from malted barley by affinity chromatography. Carlsberg Research Communications **52:** 275-283.

Brillouet J, Romieu C, Schoefs B, Solymosi K, Cheynier V, Fulcrand H, Verdeil J, Conéjéro G. 2013. The tannosome is the missing organelle forming condensed tannins of Tracheophyta. Annals of Botany **112**: 1003-1014.

Bubola M, Peršurić Đ, Ganić KK, Karoglan M, Kozina B. 2012. Effects of fruit-zone leaf removal on the concentrations of phenolic and organic acids in Istrian Malvasia grape juice and wine. Food Technology and Biotechnology **50**: 159–166.

Burbulis IE, Winkel-Shirley B. 1999. Interactions among enzymes of the Arabidopsis flavonoid biosynthetic pathway. Proceedings of the National Academy of Sciences of the United States of America **96**: 12929–12934.

Burla B, Pfrunder S, Nagy R, Francisco RM, Lee Y, Martinoia E. 2013. Vacuolar transport of abscisic acid glucosyl ester is mediated by ATP-binding cassette and proton-antiport mechanisms in Arabidopsis. Plant Physiology **163**:1446–1458.

Busse-Valverde N, Gómez-Plaza E, López-Roca JM, Gil-Muñoz R, Fernández-Fernández JI, Bautista-Ortín AB. 2010. Effect of different enological practices on skin and seed proanthocyanidins in three varietal wines. Journal of Agricultural and Food Chemistry **58**:11333-9.

Byun-McKay SA, Geeta R. 2007. Protein subcellular relocalization: a new perspective on the origin of novel genes. Trends in Ecology & Evolution **22**: 338-44.

Cadahía E, Muñoz L, Fernández de Simón B, García-Vallejo MC. 2001. Changes in low molecular weight phenolic compounds in Spanish, French, and American oak woods during natural seasoning and toasting. Journal of Agricultural and Food Chemistry **49**:1790-8.

Cadot Y, Miñana-Castelló MT, Chevalier M. 2006. Anatomical, histological, and histochemical changes in grape seeds from *Vitis vinifera* L. cv Cabernet franc during fruit development. Journal of Agricultural and Food Chemistry **54**: 9206–15.

Çakır B, Kılıçkaya O. 2013. Whole-genome survey of the putative ATP-binding cassette transporter family genes in *Vitis vinifera*. PLoS One **8**: e78860.

Calderón AA, Pedreno MA, Munoz R, Ros Barcelo A. 1993. Evidence for the non-vacuolar localization of anthocyanoplasts (anthocyanin-containing vesicles) in suspension cultured grapevine cells. Phyton **54**: 91–98.

Carbonneau A, Deloire A, Jaillard B. 2007. La vigne. Physiologie, terroir, culture. Ed 1 Vol 1. Dunod, Paris.

Carrier G, Huang YF, Le Cunff L, Fournier-Level A, Vialet S, Souquet JM, Cheynier V, Terrier N, This P. 2013. Selection of candidate genes for grape proanthocyanidin pathway by an integrative approach. Plant Physiology and Biochemistry **72**: 87–95.

Castellarin SD, Di Gaspero G, Marconi R, Nonis A, Peterlunger E, Paillard S, Adam-Blondon AF, Testolin R. 2006. Colour variation in red grapevines (*Vitis vinifera* L.): genomic organisation, expression of flavonoid 3′-hydroxylase, flavonoid 3′,5′-hydroxylase genes and related metabolite profiling of red cyanidin-/blue delphinidin-based anthocyanins in berry skin. BMC Genomics **7**: 12.

Castellarin SD, Matthews MA, Di Gaspero G, Gambetta GA. 2007. Water deficits accelerate ripening and induce changes in gene expression regulating flavonoid biosynthesis in grape berries. Planta **227**: 101–12.

Cavallini E, Matus JT, Finezzo L, Zenoni S, Loyola R, Guzzo F, Schlechter R, Ageorges A, Arce-Johnson P, Tornielli GB. 2015. The phenylpropanoid pathway is controlled at different branches by a set of R2R3-MYB C2 repressors in grapevine. Plant Physiology **167**:1448–70.

Cerezo AB, Tesfaye W, Torija MJ, Mateo E, García-Parrilla MC, Troncoso AM. 2008. The phenolic composition of red wine vinegar produced in barrels made from different woods. Food Chemistry **109**: 606–615.

Chapuis-Lardy L, Contour-Ansel D, Bernhard-Reversat F. 2002. High-performance liquid chromatography of water-soluble phenolics in leaf litter of three Eucalyptus hybrids (Congo). Plant Science **163**: 217–222.

Chen JY, Wen PF, Kong WF, Pan OH, Wan SB, Huang WD. 2006. Changes and subcellular localizations of the enzymes involved in phenylpropanoid metabolism during grape berry development. Journal of Plant Physiology **163**: 115–127.

Chen XN, Fan JF, Yue X, Wu XR, Li LT. 2008a. Radical scavenging activity and phenolic compounds in persimmon (*Diospyros kaki* L. cv. Mopan). Journal of Food Science **73**: C24–8.

Chen Z, Zhang J, Chen G. 2008b. Simultaneous determination of flavones and phenolic acids in the leaves of *Ricinus communis* Linn. by capillary electrophoresis with amperometric detection. Journal of Chromatography. B, Analytical Technology in the Biomedical Life Sciences. **863**: 101–6.

Chervin C, El-Kereamy A, Roustan JP, Latché A, Lamon J, Bouzayen M. 2004. Ethylene seems required for the berry development and ripening in grape, a non-climacteric fruit. Plant Science **167**: 1301–1305.

Cheynier V, Rigaud J. 1986. HPLC separation and characterization of flavonols in the skins of *Vitis vinifera* var Cinsault. American Journal of Enology and Viticulture **37**: 248-252.

Cheynier V, Moutounet M, Sarni-Manchado P. 1998. Les composés phénoliques. In: *Oenologie Fondements Scientifiques et Technologiques* (ed. C. Flanzy), pp. 123–162. Editions Tec & Doc, Paris.

Cheynier V. 2005. Polyphenols in foods are more complex than often thought. American Journal of Clinical Nutrition **81**: 223S–229S.

Cho MH, Corea OR, Yang H, Bedgar DL, Laskar DD, Anterola AM, Moog-Anterola FA, Hood RL, Kohalmi SE, Bernards MA, Kang C, Davin LB, Lewis NG. 2007. Phenylalanine biosynthesis in *Arabidopsis thaliana*. Identification and characterization of arogenate dehydratases. Journal of Biological Chemistry **282**: 30827–35.

Clifford MN, Knight S, Kuhnert N. 2005. Discriminating between the six isomers of dicaffeoylquinic acid by LC-MS(n). Journal of Agricultural and Food Chemistry **53**: 3821-32.

Conde C, Silva P, Fontes N, Dias ACP, Tavares RM, Sousa MJ, Agasse A, Delrot S, Geros H. 2007. Biochemical changes throughout grape berry development and fruit and wine quality. Food **1**: 1–22.

Cone KC, Burr FA, Burr B. 1986. Molecular analysis of the maize anthocyanin regulatory locus C1. Proceedings of the National Academy of Sciences **83**: 9631–5.

Cone KC, Cocciolone SM, Burr FA, Burr B. 1993. Maize anthocyanin regulatory gene *pl* is a duplicate of *c1*that functions in the plant. Plant Cell **5**: 1795–1805.

Conn S, Zhang W, Franco C. 2003. Anthocyanic vacuolar inclusions (AVIs) selectively bind acylated anthocyanins in *Vitis vinifera* L. (grapevine) suspension cultured grapevine cells. Biological Letters **25**: 835–839.

Conn S, Curtin C, Bezier A, Franco C, Zhang W. 2008. Purification, molecular cloning, and characterization of glutathione S-transferases (GSTs) from pigmented *Vitis vinifera* L. cell suspension cultures as putative anthocyanin transport proteins. Journal of Experimental Botany **59**: 3621-3634.

Conn S, Franco C, Zhang W. 2010. Characterization of anthocyanic vacuolar inclusions in *Vitis vinifera* L. cell suspension cultures. Planta **231**: 1343-60.

Constantini L, Malacarne G, Lorenzi S, Troggio M, Mattivi F, Moser C, Grando MS. 2015. New candidate genes for the fine regulation of the colour of grapes. Journal of Experimental Botany **66**: 4427-4440.

Coombe BG, Hale CR. 1973. The hormone content of ripening grape berries and the effects of growth substance treatments. Plant Physiology **51**: 629–634.

Coombe BG, McCarthy MG. 2000. Dynamics of grape berry growth and physiology of ripening. Australian Journal of Grape and Wine Research **6**: 131–135.

Cortell JM, Kennedy JA. 2006. Effect of shading on accumulation of flavonoid compounds in (*Vitis vinifera* L.) Pinot noir fruit and extraction in a model system. Journal of Agricultural and Food Chemistry **54**: 8510–8520.

Costa MA, Bedgar DL, Moinuddin SG, Kim KW, Cardenas CL, Cochrane FC, Shockey JM, Helms GL, Amakura Y, Takahashi H, Milhollan JK, Davin LB, Browse J, Lewis NG. 2005. Characterization *in vitro* and *in vivo* of the putative multigene 4-coumarate:CoA ligase network in Arabidopsis: syringyl lignin and sinapate/sinapyl alcohol derivative formation. Phytochemistry **66**: 2072–91.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator.Genome Research **14**: 1188-90.

Cutanda-Perez MC, Ageorges A, Gomez C, Vialet S, Terrier N, Romieu C, Torregrosa L. 2009. Ectopic expression of VlmybA1 in grapevine activates a narrow set of genes involved in anthocyanin synthesis and transport. Plant Molecular Biology **69**: 633–648.

Czemmel S, Stracke R, Weisshaar B, Cordon N, Harris NN, Walker AR, Robinson SP, Bogs J. 2009. The grapevine R2R3-MYB transcription factor VvMYBF1 regulates flavonol synthesis in developing grape berries. Plant Physiology **151**: 1513-30.

Czemmel S, Heppel SC, Bogs J. 2012. R2R3 MYB transcription factors: key regulators of the flavonoid biosynthetic pathway in grapevine. Protoplasma **249** (Suppl 2), S109–S118.

D'Amato TA. 1984. Subcellular localization of chorismate-mutase isoenzymes in protoplasts from mesophyll and suspension cultured cells of *Nicotiana sylvestris*. Planta **162**: 104–108.

D'Auria JC. 2006. Acyltransferases in plants: a good time to be BAHD. Current Opinion in Plant Biology **9**: 331–340.

Dal Degan F, Rocher A, Cameron-Mills V, von Wettstein D. 1994. The expression of serine carboxypeptidases during maturation and germination of the barley grain. Proceedings of the National Academy of Sciences **91**: 8209–8213.

Da Silva PPA, Laranjinha JA, de Freitas VA. 2003. Antioxidant protection of low density lipoprotein by procyanidins: structure/activity relationships. Biochemical Pharmacology **66**: 947–954.

Davies C, Robinson SP. 1996. Sugar accumulation in grape berries. Cloning of two putative vacuolar invertase cDNAs and their expression in grapevine tissues. Plant Physiology **111**:275-83.

Davies C, Boss PK, Robinson SP. 1997. Treatment of grape berries, a nonclimacteric fruit with a synthetic auxin, retards ripening and alters the expression of developmentally regulated genes. Plant Physiology **115**: 1155-1161.

Davies C, Böttcher C. 2009. Hormonal control of grape berry ripening. In Grapevine Molecular Physiology & Biotechnology, 2nd ed.; Roubelakis-Angelakis, K.A., Ed.; Springer: Berlin, Germany; pp. 229–261.

De Mello JP, Petereit F, Nahrstedt A. 1996. Prorobinetinidins from *Stryphnodendron adstringens*. Phytochemistry **42**: 857–862.

De Vasconcelos MCBM, Bennett RN, Rosa EAS, Ferreira-Cardoso JV. 2007. Primary and secondary metabolite composition of kernels from three cultivars of Portuguese chestnut (*Castanea sativa* Mill.) at different stages of industrial transformation. Journal of Agricultural and Food Chemistry **55**: 3508-3516.

Debeaujon I, Léon-Kloosterziel KM, Koornneef M. 2000. Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. Plant Physiology **122**: 403–413.

Debeaujon I, Peeters AJM, Leon-Kloosterziel KM, Koornneef M. 2001. The TRANSPARENT TESTA12 gene of Arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. Plant Cell **13**: 853–871.

Debeaujon I, Nesi N, Perez P, Devic M, Grandjean O, Caboche M, Lepiniec L. 2003. Proanthocyanidin-accumulating cells in Arabidopsis testa: regulation of differentiation and role in seed development. Plant Cell **15**: 2514–31.

Del Verde-Mendez CM, Forster MP, Rodriguez-Delgado MA, Rodriguez-Rodriguez EM, Diaz-Romero C. 2003. Content of free phenolic compounds in bananas from Tenerife (Canary Islands) and Ecuador. European Food Research and Technology **217**: 287–290.

Delcambre A, Saucier C. 2012. Identification of new flavan-3-ol monoglycosides by UHPLC-ESI-Q-TOF in grapes and wine. Journal of Mass Spectrometry **47**: 727–736.

Della-Cioppa G, Bauer SC, Klein BK, Shah DM, Fraley RT, Kishore GM. 1986. Translocation of the precursor of 5-enolpyruvylshikimate-3-phosphate synthase into chloroplasts ofhigher plants in vitro. Proceedings of the National Academy of Sciences, USA **83**: 6873–6877.

Deluc L, Barrieu F, Marchive C, Lauvergeat V, Decendit A, Richard T, Carde JP, Merillon JM, Hamdi S. 2006. Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway. Plant Physiology **140**: 499–511.

Deluc L, Bogs J, Walker AR, Ferrier T, Decendit A, Merillon JM, Robinson SP, Barrieu F. 2008. The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries. Plant Physiology **147**: 2041–2053.

Deluc LG, Quilici DR, Decendit A, Grimplet J, Wheatley MD, Schlauch KA, Mérillon JM, Cushman JC, Cramer GR. 2009. Water deficit alters differentially metabolic pathways affecting important flavor and quality traits in grape berries of Cabernet Sauvignon and Chardonnay. BMC Genomics **10**: 212.

Devic M, Guilleminot J, Debeaujon I, Bechtold N, Bensaude E, Koornneef M, Pelletier G, Delseny M. 1999. The BANYULS gene encodes a DFR-like protein and is a marker of early seed coat development. The Plant Journal **19**: 387-98.

Dewick PM, Haslam E. 1969. Phenol biosynthesis in higher plants. The Biochemical Journal **113**:537–42.

Dewick PM. 1992. The biosynthesis of shikimate metabolites.Natural Product Reports **9:** 153-181.

Diakou P, Carde JP. 2001. In situ fixation of grape berries. Protoplasma **218**: 225-235.

Diaz J, Merino F. 1997. Shikimate dehydrogenase from pepper (*Capsicum annuum*) seedlings. Purification and properties. Physiologia Plantarum **100**: 147–152**.**

Ding L, Hofius D, Hajirezaei MR, Fernie AR, Börnke F, Sonnewald U. 2007. Functional analysis of the essential bifunctional tobacco enzyme 3-dehydroquinate dehydratase/shikimate dehydrogenase in transgenic tobacco plants. Journal of Experimental Botany **58:** 2053–2067.

Dixon RA, Paiva NL. 1995. Stress-induced phenylpropanoid metabolism. Plant Cell **7**:1085–1097.

Dixon RA, Xie DY, Sharma SB. 2005. Proanthocyanidins - a final frontier in flavonoid research? New Phytologist **165**: 9-28.

Doan NP, Fincher GB. 1988. The A- and B-chains of carboxypeptidase I from germinated barley originate from a single precursor polypeptide. Journal of Biological Chemistry **263**: 11106–11110.

Dóminguez F, González MC, Cejudo FJ. 2002. A germination-related gene encoding a serine carboxypeptidase is expressed during the differentiation of the vascular tissue in wheat grains and seedlings. Planta **215**: 727-34.

Downey MO, Harvey JS, Robinson SP. 2003a. Synthesis of flavonols and expression of flavonol synthase genes in the developing grape berries of Shiraz and Chardonnay (*Vitis vinifera* L.). Australian Journal of Grape and Wine Research **9**: 110–121.

Downey MO, Harvey JS, Robinson SP. 2003b. Analysis of tannins in seeds and skins of Shiraz grapes throughout berry development. Australian Journal of Grape and Wine Research **9**: 15–27.

Downey MO, Harvey JS, Robinson SP. 2004. The effect of bunch shading on berry development and flavonoid accumulation in Shiraz grapes. Australian Journal of Grape and Wine Research **1**: 55–73.

Dubos C, Le Gourrierec J, Baudry A, Huep G, Lanet E, Debeaujon I, Routaboul JM, Alboresi A, Weisshaar B, Lepiniec L. 2008. MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. Plant Journal **55**: 940–953.

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. 2010. MYB transcription factors in Arabidopsis. Trends in Plant Science **15**: 573–81.

Dudareva N, D'Auria JC, Nam KH, Raguso RA, Pichersky E. 1998. Acetyl-CoA:benzylalcohol acetyltransferase - an enzyme involved in floral scent production in *Clarkia breweri*. *The Plant Journal* **14**: 297–304.

Duke SO, Powles SB. 2008. Glyphosate: a once-in-a-century herbicide. Pest Management Science **64**: 319-25.

Ebadi A, Coombe BG, May P. 1995. Fruit-set on small Chardonnay and Shiraz vines grown under varying temperature regimes between budburst and flowering. Australian Journal of Grape and Wine Research **1**: 3–10.

Egli DB. 1998. Seed biology and the yield of grain crops. Plant Growth Regulation **26**:140-141.

El-Basyouni SZ, Chen D, Ibrahim RK, Neish AC, Towers GHN. 1964. The biosynthesis of hydroxybenzoic acids in higher plants.Phytochemistry **3**: 485–492.

El-Seedi HR, El-Said AM, Khalifa SA, Göransson U, Bohlin L, Borg-Karlson AK, Verpoorte R. 2012. Biosynthesis, natural sources, dietary intake, pharmacokinetic properties, and biological activities of hydroxycinnamic acids. Journal of Agricultural and Food Chemistry **60**:10877-95.

Eijkman JF. 1885. Sur les principes constituants de l'*Illicium religiosum* (Sieb.) (Shikimi-no-ki en japonais). Recueil des Travaux Chimiques des Pays-Bas **4**: 32–54.

El-Kereamy A, Chervin C, Roustan JP, Cheynier V, Souquet JM, Moutounet M, Raynal J, Ford CM, Latche A, Pech JC, Bouzayen M. 2003. Exogenous ethylene stimulates the longterm expression of genes related to anthocyanin biosynthesis in grape berries. Plant physiology **119**: 175-182.

Emanuelsson O, Nielsen H, von Heijne G. 2006. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Science **8**: 978–984.

Endrizzi JA, Breddam K, Remington SJ. 1994. 2.8-A structure of yeast serine carboxypeptidase. Biochemistry **33**: 11106-20.

Entus R, Poling M, Herrmann KM. 2002. Redox regulation of Arabidopsis 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. Plant Physiology **129**: 1866–1871.

Es-Safi NE, Guyot S, Ducrot PH. 2006. NMR, ESI/MS, and MALDI-TOF/MS analysis of pear juice polymeric proanthocyanidins with potent free radical scavenging activity. Journal of Agricultural and Food Chemistry **54**:6969-77.

Esmaeili N, Ebrahimzadeh H, Abdi K, Safarian S. 2011. Determination of some phenolic compounds in *Crocus sativus* L. corms and its antioxidant activities study. Pharmacognosy Magazine **7**:74-80.

Eudes A, Bozzo GG, Waller JC, Naponelli V, Lim EK, Bowles DJ, Gregory JF 3rd, Hanson AD. 2008. Metabolism of the folate precursor p-aminobenzoate in plants: glucose ester formation and vacuolar storage. The Journalof Biological Chemistry **283**: 15451–9.

Fabre S, Pinaud N, Fouquet E, Pianet I. 2010. Colloidal behavior of wine galloylated tannins. Comptes Rendus Chimie **13:** 561–565.

Falcone Ferreyra ML, Rius SP, Casati P. 2012. Flavonoids: biosynthesis, biological functions, and biotechnological applications. Frontiers in Plant Science **3**:222.

Falginella L, Castellarin SD, Testolin R, Gambetta GA, Morgante M, Di Gaspero G. 2010. Expansion and subfunctionalisation of flavonoid 3',5'-hydroxylases in the grapevine lineage. BMC Genomics **11**: 562.

Fang Z, Zhang M, Wang L. 2007. HPLC-DAD-ESIMS analysis of phenolic compounds in bayberries (*Myrica rubra* Sieb. et Zucc.). Food Chemistry **100**: 845–852.

Fang Z, Hu Y, Liu D, Chen J, Ye X. 2008. Changes of phenolic acids and antioxidant activities during potherb mustard (*Brassica juncea*, Coss.) pickling. Food Chemistry **108**: 811–7.

Fang F, Tang K, Huang WD. 2013. Changes of flavonol synthase and flavonol contents during grape berry development. European Food Research and Technology **237**: 529-540.

Fazio G, Gattuso A, Cilluffo V, Arcoleo G. 1983. Preparation and characterization of protein materials from grape seed meals. Rivista della Societa Italiana di Scienza dell'Alimentazione **6**: 469–478.

Feng H, Li Y, Wang S, Zhang L, Liu Y, Xue F, Sun Y, Wang Y, Sun J. 2014. Molecular analysis of proanthocyanidins related to pigmentation in brown cotton fibre (*Gossypium hirsutum* L.). Journal of Experimental Botany **65**: 5759–69.

Fernández de Simón B, Sanz M, Cadahía E, Poveda P, Broto M. 2006. Chemical characterization of oak heartwood from Spanish forests of *Quercus pyrenaica* (Wild.). Ellagitannins, low molecular weight phenolic, and volatile compounds. Journal of Agricultural and Food Chemistry **54**: 8314–21.

Ferrandino A, Carra A, Rolle L, Schneider A, Schubert A. 2012. Profiling of hydroxycinnamoyl tartrates and acylated anthocyanins in the skin of 34 *Vitis vinifera* genotypes. Journal of Agricultural and Food Chemistry. **60**: 4931-45.

Ferrari S, Galletti R, Denoux C, Lorenzo G, Ausubel F, Dewdney J. 2007. Resistance to *Botrytis cinerea* induced in Arabidopsis by elicitors is independent of salicylic acid, ethylene, or jasmonate signaling but requires PHYTOALEXIN DEFICIENT3. Plant Physiology **144**: 367–379.

Feuereisen MM, Hoppe J, Zimmermann BF, Weber F, Schulze-Kaysers N, Schieber A. 2014. Characterization of phenolic compounds in Brazilian pepper (*Schinus terebinthifolius* Raddi) exocarp. Journal of Agricultural and Food Chemistry **62**:6219-26.

Fiedler E, Schultz G. 1985. Localization, purification, and characterization of shikimate oxidoreductase-dehydroquinate hydrolyase from stroma of spinach chloroplasts. Plant Physiology **79**: 212–218.

Flagel LE, Wendel JF. 2009. Gene duplication and evolutionary novelty in plants. New Phytologist **183**: 557-64.

Flamini R, Mattivi F, De Rosso M, Arapitsas P, Bavaresco L. 2013. Advanced knowledge of three important classes of grape phenolics: anthocyanins, stilbenes and flavonols. International Journal of Molecular Science. **14**: 19651–69.

Ford CM, Boss PK, Hoj PB. 1998. Cloning and characterization of *Vitis vinifera* UDP-glucose:flavonoid 3-*O*-glucosyltransferase, a homologue of the enzyme encoded by the maize Bronze-1 locus that may primarily serve to glucosylate anthocyanidins in vivo. Journal of Biological Chemistry **273**: 9224-33.

Forkmann G, Heller W, Grisebach H. 1980. Anthocyanin biosynthesis in flowers of Matthiola-Incana flavanone-3-hydroxylase and flavonoid-3'-hydroxylase. Zeitschrift Fur Naturforschung C - A Journal of Biosciences **35**: 691-695.

Fornara V, Onelli E, Sparvoli F, Rossoni M, Aina R, Marino G, Citterio S. 2008. Localization of stilbene synthase in *Vitis vinifera* L. during berry development. Protoplasma **233**:83-93.

Fortes AM, Teixeira RT, Agudelo-Romero P. 2015. Complex interplay of hormonal signals during grape berry ripening. Molecules **20**:9326-43.

Fossen T, Larsen Å, Andersen ØM. 1998. Anthocyanins from flowers and leaves of *Nymphaéa* x *marliacea* cultivars. Phytochemistry **48**: 823–827.

Fossen T, Andersen ØM. 1999a. Delphinidin 3'-galloylgalactosides from blue flowers of *Nymphaéa caerulea*. Phytochemistry **50**: 1185–1188.

Fossen T, Andersen ØM. 1999b. Cyanidin 3-(2",3"-digalloylglucoside) from red leaves of Acer. Phytochemistry **52**: 1697–1700.

Fournand D, Vicens A, Sidhoum L, Souquet JM, Moutounet M, Cheynier V. 2006. Accumulation and extractability of grape skin tannins and anthocyanins at different advanced physiological stages. Journal of Agricultural and Food Chemistry **54**: 7331-7338.

Fournand D, Moutounet M. 2008. La maturité phénolique du raisin : un concept en cours de maturation. Revue Française d'Œnologie **228**: 1-6.

Francisco RM, Regalado A, Ageorges A, Burla BJ, Bassin B, Eisenach C, Zarrouk O, Vialet S, Marlin T, Chaves MM, Martinoia E, Nagya R. 2013. ABCC1, an ATP binding cassette protein from grape berry, transports anthocyanidin 3-*O*-glucosides. The Plant Cell **25**: 1840–1854.

Frangne N, Eggmann T, Koblischke C, Weissenbock G, Martinoia E, Klein M. 2002. Flavone glucoside uptake into barley mesophyll and arabidopsis cell culture vacuoles. Energization occurs by H+-antiport and ATP-binding cassette-type mechanisms. Plant Physiology **128**: 726-733.

Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C. 2002. The Arabidopsis REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. Plant Journal **30**: 33-45.

Fraser CM, Rider LW, Chapple C. 2005. An expression and bioinformatics analysis of the Arabidopsis serine carboxypeptidase like gene family. Plant Physiology **138**: 1136–1148.

Fraser CM, Thompson MG, Shirley AM, Ralph J, Schoenherr JA, Sinlapadech T, Hall MC, Chapple C. 2007. Related Arabidopsis serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. Plant Physiology **144**: 1986–1999.

Froemel S, Devlaming P, Stotz G, Wiering H, Forkmann G, Schram AW. 1985. Genetic and biochemical studies on the conversion of flavanones to dihydroflavonols in flowers of *Petunia hybrida*. Theoretical and Applied Genetics **70**: 561–568.

Fröhlich B, Niemetz R, Gross GG. 2002. Gallotannin biosynthesis: two new galloyltransferases from *Rhus typhina* leaves preferentially acylating hexa- and heptagalloylglucoses. Planta **216**: 168-72.

Fucile G, Falconer S, Christendat D. 2008. Evolutionary diversification of plant shikimate kinase gene duplicates. PLoS Genet. **4**, e1000292.

Fujiwara H, Tanaka Y, Yonekura-Sakakibara K, Fukuchi-Mizutani M, Nakao M, Fukui Y, Yamaguchi M, Ashikari T, Kusumi T. 1998. cDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from *Gentiana triflora*. The Plant Journal **16**: 421–431.

Fung RWM, Gonzalo M, Fekete C, Kovacs LG, He Y, Marsh E, McIntyre LM, Schachtman DP, Qiu W. 2008. Powdery mildew induces defense-oriented reprogramming of the transcriptome in a susceptible but not in a resistant grapevine. Plant Physiology **146**:236–249.

Gagné S, Saucier C, Gény L. 2006. Composition and cellular localization of tannins in Cabernet sauvignon skins during growth. Journal of Agriculture and Food Chemistry **54**: 9465-9471.

Galzy R. 1964. Technique de thermotherapie des viroses de la vigne. Annales des Epiphyties **15**:245–256.

Ganson RJ, D'Amato TA, Jensen RA. 1986. The two-isozyme system of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase in *Nicotiana silvestris* and other higher plants. Plant Physiology **82**: 203–210.

Gambetta GA, Matthews MA, Shaghasi TH, McElrone AJ, Castellarin SD. 2010. Sugar and abscisic acid signalling orthologs are activated at the onset of ripening in grape. Planta **232**: 219-34.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research **31**: 3784–8.

Gatto P, Vrhovsek U, Muth J. 2008. Ripening and genotype control stilbene accumulation in healthy grape. Journal of Agricultural and Food Chemistry **56**: 11773–11785.

Gawel R, Oberholster A, Francis IL. 2000. The "Mouth-feel Wheel": terminology for communicating the mouth-feel characteristics of red wine. Australian Journal of Grape and WineResearch **6**: 203–207.

Gény L, Saucier C, Bracco S, Daviaud F, Glories Y. 2003. Composition and cellular localization of tannins in grape seeds during maturation. Journal of Agricultural and Food Chemistry **51**: 8051–4.

Glories Y. 1998. Les composés phénoliques. La notion de maturation phénolique. In 'Traité d'œnologie. Volume II. Chimie du vinstabilisation et traitements'. (Ed. P Ribereau-Gayon) pp. 220-221. (Dunod: Paris).

Gomez C, Terrier N, Torregrosa L, Vialet S, Fournier-Level A, Verriès C, Souquet JM, Mazauric JP, Klein M, Cheynier V, Ageorges A. 2009. Grapevine MATE-type proteins act as vacuolar H^+-dependent acylated anthocyanin transporters. Plant Physiology **150**: 402–415.

Gomez C, Conejero G, Torregrosa L, Cheynier V, Terrier N, Ageorges A. 2011. In vivo grapevine anthocyanin transport involves vesicle mediated trafficking and the contribution of anthoMATE transporters and GST. The Plant Journal **67**: 960–970.

Gomord V, Fitchette AC, Menu-Bouaouiche L, Saint-Jore-Dupas C, Plasson C, Michaud D, Faye L. 2010. Plant-specific glycosylation patterns in the context of therapeutic protein production. Plant Biotechnology Journal **8**: 564–87.

Goodman CD, Casati P, Walbot V. 2004. A multidrug resistance-associated protein involved in anthocyanin transport in *Zea mays*. Plant Cell **16**: 1812-1826.

Gordon A, Friedrich M, da Matta VM, Herbster Moura CF, Marx F. 2012. Changes in phenolic composition, ascorbic acid and antioxidant capacity in cashew apple (*Anacardium occidentale* L.) during ripening. Fruits **67**: 267–276.

Gorinstein S, Cvikrova M, Machackova I, Ratiporn H, Yong Seo P, Soon Teck J, Yamamoto K, Martinez-Ayala AL, Katrich E, Trakhtenberg S. 2004. Characterization of antioxidant compounds in Jaffa sweeties and white grapefruits. Food Chemistry **84**: 503–510.

Goto-Yamamoto N, Wang GH, Masaki K, Kobayashi S. 2002. Structure and transcription of three chalcone synthase genes of grapevine (*Vitis vinifera*). Plant Science **162**: 867–872.

Greenspan MD, Shackel KA, Matthews MA. 1994. Developmental changes in the diurnal water budget of the grape berry exposed to water deficits. Plant, Cell and Environment **17**: 811–820.

Grotewold E. 2004. The challenges of moving chemicals within and out of cells: insights into the transport of plant natural products. Planta **219**: 906–909.

Grotewold E. 2006. The genetics and biochemistry of floral pigments. Annual Review of Plant Biology **57**: 761–780.

Guo J, Carrington Y, Alber A, Ehlting J. 2014. Molecular characterization of quinate and shikimate metabolism in *Populus trichocarpa*. Journal of Biological Chemistry **289**:23846-58.

Hamberger B, Ehlting J, Barbazuk B, Douglas CJ. 2006. Comparative genomics of the shikimate pathway in *arabidopsis*, *populus trichocarpa* and *oryza sativa*: Shikimate pathway gene family structure and identification of candidates for missing links in phenylalanine biosynthesis. Recent Advances in Phytochemistry **40**: 85–113.

Hancock K R, Collette V, Fraser K, Greig M, Xue H, Richardson K, Jones C, Rasmussen S. 2012. Expression of the R2R3-MYB transcription factor TaMYB14 from *Trifolium arvense* activates proanthocyanidin biosynthesis in the legumes *Trifolium repens* and *Medicago sativa*. Plant Physiology **159**: 1204–1220.

Harbone JB, Williams CA. 2000. Advances in flavonoid research since 1992. Phytochemistry **55**: 481-504.

Hardie WJ, Aggenbach SJ. 1996. Effects of site, season and viticultural practices on grape seed development. Australian Journal of Grape and Wine Research **2**: 21-24.

Hardie WJ, O'Brien TP, Jaudzems VG. 1996. Morphology, anatomy and development of the pericarp after anthesis in grape, *Vitis vinifera* L. Australian Journal of Grape Wine Research **2**:97-142.

Harris JM, Kriedemann PE, Possingham JV. 1968. Anatomical aspects of grape berry development. Vitis **7**: 106-119.

Haslam E. 1993. Shikimic acid, metabolism and metabolites. Chichester: John Wiley and Sons.

Hause B, Meyer K, Viitanen PV, Chapple C, Strack D. 2002. Immunolocalization of 1-*O*-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. Planta **215**:26–32.

Heller W, Forkmann G. 1988. Biosynthesis. In: The Flavonoids – Advances in Research since 1980 (ed. J.B. Harborne), pp. 399–425. Chapman & Hall, London.

Herrmann KM. 1995. The shikimate pathway: early steps in the biosynthesis of aromatic compounds. The Plant Cell **7:** 907-919.

Hernández I, Alegre L, Munné-Bosch S. 2006. Enhanced oxidation of flavan-3-ols and proanthocyanidin accumulation in water-stressed tea plants. Phytochemistry **67**:1120-6.

Herrmann KM. 1995. The shikimate pathway: early steps in the biosynthesis of aromatic compounds. Plant Cell **7**:907-919.

Herrmann KM, Weaver LM. 1999. The shikimate pathway. Annual Review of Plant Physiology and Plant Molecular Biology **50**:473–503.

Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V. 2011. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. Journal of Experimental Botany **62**: 2465–2483.

Himi E, Yamashita Y, Haruyama N, Yanagisawa T, Maekawa M, Taketa S. 2012. Ant28 gene for proanthocyanidin synthesis encoding the R2R3 MYB domain protein (Hvmyb10) highly affects grain dormancy in barley. Euphytica **188**: 141-151.

Hoang NTT, Golding JB, Wilkes MA. 2011. The effect of postharvest 1-MCP treatment and storage atmosphere on 'Cripps Pink' apple phenolics and antioxidant activity. Journal of Agricultural and Food Chemistry **55**: 8616–24.

Hoffmann L, Maury S, Martz F, Geoffroy P, Legrand M. 2003. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. The Journal of Biological Chemistry **278**: 95–103.

Holton TA, Brugliera F, Tanaka Y. 1993. Cloning and expression of flavonol synthase from *Petunia hybrida*. Plant Journal **4**: 1003-1010.

Hrazdina G, Zobel AM, Hoch HC. 1987. Biochemical, immunological, and immunocytochemical evidence for the association of chalcone synthase with endoplasmic reticulum membranes. Proceedings of the National Academy of Sciences **84**: 8966–70.

Hrazdina G, Jensen RA. 1992. Spatial-Organization of Enzymes in Plant Metabolic pathways. Plant Physiology and Plant Molecular Biology **43**: 241-267.

Hu Z, Shen Y, Shen F, Luo Y, Su X. 2009. Evidence for the signaling role of methyl jasmonate, methyl salicylate and benzothiazole between poplar (*Populus simonii x P. pyramidalis* 'Opera 8277') cuttings. Trees **23**:1003–1011.

Huang T, Tohge T, Lytovchenko A, Fernie AR, Jander G. 2010. Pleiotropic physiological consequences of feedback-insensitive phenylalanine biosynthesis in *Arabidopsis thaliana*. Plant Journal **63**: 823–35.

Huang YF, Cheynier V, Terrier N. 2012. Shedding light on the black boxes of the proanthocyanidin pathway with grapevine. In: Veronique Cheynier, Pascale Sarni-Manchado, Stephane Quideau, Recent Advances in Polyphenol Research (p. 161-190). Recent Advances in Polyphenol Research, 3. Malden, USA : Wiley-Blackwell.

Huang YF, Doligez A, Fournier-Level A, Le Cunff L, Bertrand Y, Canaguier A, Morel C, Miralles V, Veran F, Souquet JM, Cheynier V, Terrier N, This P. 2012. Dissecting genetic architecture of grape proanthocyanidin composition through quantitative trait locus mapping.BMC Plant Biology **12**:30.

Huang YF, Vialet S, Guiraud JL, Torregrosa L, Bertrand Y, Cheynier V, This P, Terrier N. 2014. A negative MYB regulator of proanthocyanidin accumulation, identified through expression quantitative locus mapping in the grape berry. New Phytologist **201**: 795–809.

Hugueney P, Provenzano S, Verriès C, Ferrandino A, Meudec E, Batelli G, Merdinoglu D, Cheynier V, Schubert A, Ageorges A. 2009. A novel cation-dependent O-methyltransferase involved in anthocyanin methylation in grapevine. Plant Physiology **150**: 2057-70.

Humphreys JM, Hemm MR, Chapple C. 1999. New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proceedings of the National Academy of Sciences **96**: 10045–50.

Igartuburu JM, Martín del Río R, Massanet GM, Montiel JA, Pando E, Rodríguez Luis F. 1991. Study of agricultural by-products. Extractability and amino acid composition of grapeseed (*Vitis vinifera*) proteins. Journal of the Science of Food and Agriculture **54**: 489–493.

Ikegami A, Sato A, Yamada M, Kitajima A, Yonemori K. 2005. Expression of genes involved in proanthocyanidin biosynthesis during fruit development in a Chinese pollination-constant, nonastringent (PCNA) persimmon, 'Luo Tian Tian Shi'. Journal of American Society of Horticultural Science **130:** 830–835.

Ikegami A, Eguchi S, Kitajima A, Inoue K, Yonemori K. 2007. Identification of genes involved in proanthocyanidin biosynthesis of persimmon (*Diospyros kaki*) fruit. Plant Science **172**: 1037–1047.

Inoue M, Sakaguchi N, Isuzugawa K, Tani H, Ogihara Y. 2000. Role of reactive oxygen species in gallic acid-induced apoptosis.Biological and Pharmaceutical Bulletin **23**:1153-7.

Ishikura N, Hayashida S, Tazaki K. 1984. Biosynthesis of gallic and ellagic acids with^{14}C-labeled compounds in*Acer* and*Rhus* leaves.The Botanical Magazine **97**: 355-367.

Jaillon O, Aury JM, Noël B. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature **449**: 463-7.

Jean-Denis JB, Pezet R, Tabacchi R. 2006. Rapid analysis of stilbenes and derivatives from downy mildew-infected grapevine leaves by liquid chromatography–atmospheric pressure photoionisation mass spectrometry. Journal of Chromatography A **1112**: 263–8.

Jeong ST, Goto-Yamamoto N, Kobayashi S, Esaka M. 2004. Effects of plant hormones and shading on the accumulation of anthocyanins and the expression of anthocyanin biosynthetic genes in grape berry skins. Plant Science **167**: 247–252.

Jeong ST, Goto-Yamamoto N, Hashizume K, Esaka M. 2006. Expression of the flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase genes and flavonoid composition in grape (*Vitis vinifera*). Plant Science **170**: 61–69.

Ji SB, Saito N, Yokoi M, Shigihara A, Honda T. 1992. Galloylcyanidin glycosides from acer. Phytochemistry **31**: 655–657.

Jiang X, Liu Y, Li W, Zhao L, Meng F, Wang Y, Tan H, Yang H, Wei C, Wan X, Gao L, Xia T. 2013. Tissue-specific, development-dependent phenolic compounds accumulation profile and gene expression pattern in tea plant [*Camellia sinensis*]. PLoS One **8:** e62315.

Kambourakis S, Frost JW. 2000. Synthesis of gallic acid: Cu^{2+}-mediated oxidation of 3-dehydroshikimic acid. Journal of Organic Chemistry **65**: 6904-6909.

Karamat F, Olry A, Munakata R, Koeduka T, Sugiyama A, Paris C, Hehn A, Bourgaud F, Yazaki K. 2014. A coumarin-specific prenyltransferase catalyzes the crucial biosynthetic reaction for furanocoumarin formation in parsley. Plant Journal **77**: 627–38.

Karawajczyk A, Drgan V, Medic N, Oboh G, Passamonti S, Novic M. 2007. Properties of flavonoids influencing the binding to bilitranslocase investigated by neural network modelling. Biochemical Pharmacology **73**: 308–20.

Karimi M, Inze D, Depicker A. 2002. GATEWAY((TM)) vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science **7**: 193–195.

Kasai K, Kanno T, Akita M, Ikejiri-Kanno Y, Wakasa K, Tozawa Y. 2005. Identification of three shikimate kinase genes in rice: characterization of their differential expression during panicle development and of the enzymatic activities of the encoded proteins. Planta **222**: 438–447.

Kato N, Shiroya M, Yoshida S, Hasegawa M. 1968. Biosynthesis of gallic acid by a homogenate of the leaves of *Pelargonium inquinans*. The Botanical Magazine **81**: 506.

Keeling CI, Weisshaar S, Lin RPC, Bohlmann J. 2008. Functional plasticity of paralogous diterpene synthases involved in conifer defense. Proceedings of the National Academy of Sciences, USA **105**: 1085–1090.

Keller M. 2015. The Science of Grapevines: Anatomy and Physiology. Second Edition. Academic Press, Oxford.

Kennedy JA, Troup GJ, Pilbrow JR, Hutton DR, Hewitt D, Hunter CR,Ristic R, Iland PG, Jones GP. 2000. Development of seed polyphenols in berries from *Vitisvinifera* L. cv. Shiraz. Australian Journal of Grape and Wine Research **6**: 244–254.

Kennedy JA, Hayasaka Y, Vidal S, Waters EJ, Jones GP. 2001. Composition of grape skin proanthocyanidins at different stages of berry development. Journal of Agricultural and Food Chemistry **49**: 5348–5355.

Kennedy JA, Matthews MA, Waterhouse AL. 2002. Effect of maturity and vine water status on grape skin and wine flavonoids. American Journal of Enology and Viticulture **53**: 268–274.

Kennedy JA, Saucier C, Glories Y. 2006. Grape and wine phenolics: History and perspective. American Journal of Enology and Viticulture **57**: 239-248.

Khallouki F, Haubner R, Erben G, Ulrich CM, Owen RW. 2012. Phytochemical composition and antioxidant capacity of various botanical parts of the fruits of *Prunus* × *domestica* L. from the Lorraine region of Europe. Food Chemistry **133**: 697–706.

Khater Fida. "Identification et validation fonctionnelle de nouveaux genes potentiellement implqués dans la biosynthèse des composés phénoliques". Thèse de doctorat en biologie, Montpellier SupAgro, 2011, 203 p.

Khater F, Fournand D, Vialet S, Meudec E, Cheynier V, Terrier N. 2012. Identification and functional characterization of cDNAs coding for hydroxybenzoate/hydroxycinnamate glucosyltransferases co-expressed with genes related to proanthocyanidin biosynthesis. Journal of Experimental Botany **63**: 1201–1214.

Kim HJ, Woo ER, Shin CG, Park H. 1998. A new flavonol glycoside gallate ester from *Acer okamotoanum* and its inhibitory activity against Human Immunodeficiency Virus-1 (HIV-1) Integrase. Journal of Natural Products **61**: 145–148.

Kingston-Smith AH. 2001. Resource allocation. Trends in Plant Science **6**: 48–49.

Kitamura S, Shikazono N, Tanaka A. 2004. TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in Arabidopsis. The Plant Journal **37**: 104–114.

Kitamura S. 2006. Transport of flavonoids: from cytosolic synthesis to vacuolar accumulation. In: Grotewold E, editor. Science of flavonoids. Berlin, Germany: Springer 123–146.

Kitamura S, Matsuda F, Tohge T, Yonekura-Sakakibara K, Yamazaki M, Saito K, Narumi I. 2010. Metabolic profiling and cytological analysis of proanthocyanidins in immature seeds of *Arabidopsis thaliana* flavonoid accumulation mutants. Plant Journal **62**: 549–559.

Kivilompolo M, Oburka V, Hyotylainen T. 2007. Comparison of GC-MS and LC-MS methods for the analysis of antioxidant phenolic acids in herbs. Analytical and Bioanalytical Chemistry **388**: 881–887.

Kobayashi S, Ishimaru M, Hiraoka K, Honda C. 2002. Myb-related genes of the Kyoho grape (*Vitis labruscana*) regulate anthocyanin biosynthesis. Planta **215**: 924–33.

Kobayashi S, Goto-Yamamoto N, Hirochika H. 2004. Retrotransposon-induced mutations in grape skin color. Science **304**: 982.

Koch R, Alleweldt G. 1978. Der Gaswechselreifender Weinbeeren.Vitis **17**: 30–44.

Koes RE, Quattrocchio F, Mol JNM. 1994. The flavonoid biosynthetic pathway in plants -function and evolution. Bioessays **16**: 123–132.

Koes R, Verweij W, Quattrocchio F. 2005. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends Plant Science **10**: 236–242.

Koornneef M. 1990. Mutations affecting the testa colour in Arabidopsis. Arabidopsis Information Service **27**: 1–4.

Koshiba T. 1979. Alicyclic acid metabolism in plants 12. Partial purification and some properties of shikimate kinase from *Phaseolus mungo* seedlings. Plant Cell Physiology **20**: 803–809.

Koukol J, Conn EE. 1961. Metabolism of aromatic compounds in higher plants. 4. Purification and properties of phenylalanine deaminase of *Hordeum vulgare*. Journal of Biological Chemistry **236**: 2692–2710.

Kowalski R, Kowalska G. 2005. Phenolic acid contents in fruits of aubergine (*Solanum Melongena* L.). Polish Journal of Food and Nutrition Sciences **14/55**:37–42.

Koyama K, Numata M, Nakajima I, Goto-Yamamoto N, Matsumura H, Tanaka N. 2014. Functional characterization of a new grapevine MYB transcription factor and regulation of proanthocyanidin biosynthesis in grapes. Journal of Experimental Botany **65**: 4433–49.

Krenek KA, Barnes RC, Talcott ST. 2014. Phytochemical composition and effects of commercial enzymes on the hydrolysis of gallic acid glycosides in mango (*Mangifera indica* L. cv. 'Keitt') pulp. Journal of Agricultural and Food Chemistry **62**:9515–21.

Kreuzaler F, Hahlbrock K. 1972. Enzymatic synthesis of aromatic compounds in higher plants: formation of naringenin (5,7,4'-trihydroxyflavanone) from *p*-coumaroyl coenzyme A and malonyl coenzyme A. FEBS Letters **15**: 69–72.

Kruger MJ, Davies N, Myburgh KH, Lecour S. 2014. Proanthocyanidins, anthocyanins and cardiovascular diseases.Food Research International **59**: 41–52.

Kulich I, Žárský V. 2014. Autophagy-related direct membrane import from ER/cytoplasm into the vacuole or apoplast: a hidden gateway also for secondary metabolites and phytohormones ? International Journal of Molecular Science **15**: 7462-7474.

Lacampagne S, Gagné S, Gény L. 2010. Involvement of abscisic acid in controlling the proanthocyanidin biosynthesis pathway in grape skin: New elements regarding the regulation of tannin composition and leucoanthocyanidin reductase (LAR) and anthocyanidin reductase (ANR) activities and expression. Journal of Plant Growth Regulation **29**: 81–90.

Lai Y, Li H, Yamagishi M. 2013. A review of target gene specificity of flavonoid R2R3-MYB transcription factors and a discussion of factors contributing to the target gene selectivity. Frontiers in Biology **8**: 577-598

Lallemand LA, Zubieta C, Lee SG, Wang Y, Acajjaoui S, Timmins J, McSweeney S, Jez JM, McCarthy JG, McCarthy AA. 2012. A structural basis for the biosynthesis of the major chlorogenic acids found in coffee. Plant Physiology **160**: 249-60.

Lambert M, Meudec E, Verbaere A, Mazerolles G, Wirth J, Masson G, Cheynier V, Sommerer N. 2015. A High-Throughput UHPLC-QqQ-MS Method for Polyphenol Profiling in Rosé Wines. Molecules **20**: 7890–914.

Landete JM. 2011. Ellagitannins, ellagic acid and their derived metabolites: A review about source, metabolism, functions and health. Food Research International **44**: 1150–1160.

Lee VS, Dou J, Chen RJ, Lin RS, Lee MR, Tzen JT. 2008. Massive accumulation of gallic acid and unique occurrence of myricetin, quercetin, and kaempferol in preparing old oolong tea.Journal of Agricultural and Food Chemistry **56:**7950–6.

Lee S, Kaminaga Y, Cooper B, Pichersky E, Dudareva N, Chapple C. 2012. Benzoylation and sinapoylation of glucosinolate R-groups in Arabidopsis. The Plant Journal **72**: 411–422.

Lehfeldt C, Shirley AM, Meyer K, Ruegger MO, Cusumano JC, Viitanen PV, Strack D, Chapple C. 2000. Cloning of the SNG1 gene of Arabidopsis reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. The Plant Cell **12**: 1295–1306.

Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M. 2006. Genetics and biochemistry of seed flavonoids. Annual Review of Plant Biology **57**: 405-430.

Leuschner C, Schultz G. 1991. Uptake of shikimate pathway intermediates by intact chloroplasts. Phytochemistry **30**: 2203-2207.

Lewin B. 1994. Genes V. Oxford University Press, New York.

Li AX, Eannetta N, Ghangas GS, Steffens JC. 1999. Glucose polyester biosynthesis. Purification and characterization of a glucose acyltransferase. Plant Physiology **121**: 453–460.

Li AX, Steffens JC. 2000. An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. Proceedings of the National Academy of Sciences of the USA **97**: 6902–6907.

Li C, Leverence R, Trombley JD, Xu S, Yang J, Tian Y, Reed JD, Hagerman AE. 2010. High molecular weight persimmon (*Diospyros kaki* L.) proanthocyanidin: a highly galloylated, A-linked tannin with an unusual flavonol terminal unit, myricetin. Journal of Agricultural and Food Chemistry **58**: 9033-42.

Li J, Lease KA, Tax FE, Walker JC. 2001. BRS1, a serine carboxypeptidase, regulates BRI1 signaling in Arabidopsis thaliana. Proceedings of the National Academy of Sciences **98**: 5916-21.

Li K, Mikola MR, Draths KM, Worden RM, Frost JW. 1999. Fedbatch fermentor synthesis of 3-dehydroshikimic acid using recombinant *Escherichia coli*. Biotechnology and Bioengineering **64:**61–73.

Li P, Wang XQ, Wang HZ, Wu YN. 1993. High performance liquid chromatographic determination of phenolic acids in fruits and vegetables. Biomedical and Environmental Sciences **6**: 389–398.

Liao DI, Remington SJ. 1990. Structure of wheat serine carboxypeptidase II at 3.5-A resolution. A new class of serine proteinase. Journal of Biological Chemistry **265**: 6528-31.

Liu RN, Wang W, Ding Y, Xie WD, Ma C, Du LJ. 2007. A new flavonol glycoside and activity of compounds from the flower of *Nymphaea candida*. Journal of Asian Natural Products Research **9**: 333-8.

Liu Y, Wu Y, Yuan K, Ji C, Hou A, Yoshida T, Okuda T. 1997. Astragalin 2″,6″-di-*O*-gallate from *Loropetalum chinense*. Phytochemistry **46**: 389–391.

Liu Y, Gao L, Xia T, Zhao L. 2009. Investigation of the site-specific accumulation of catechins in the tea plant [*Camellia sinensis* (L.) O. Kuntze] via vanillin-HCl staining. Journal of Agricultural and Food Chemistry **57**: 10371–10376.

Liu D, Shi L, Han C, Yu J, Li D, Zhang Y. 2012. Validation of reference genes for gene expression studies in virus-infected *Nicotiana benthamiana* using quantitative real-time PCR. PLoS One **7**: e46451.

Liu Y, Gao L, Liu L, Yang Q, Lu Z, Nie Z, Wang Y, Xia T. 2012. Purification and characterization of a novel galloyltransferase involved in catechin galloylation in the tea plant (*Camellia sinensis*). The Journal of Biological Chemistry **287**: 44406–44417.

Liu J, Osbourn A, Ma P. 2015a. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. Molecular Plant **8**: 689–708.

Liu Y, Shi Z, Maximova SN, Payne MJ, Guiltinan MJ. 2015b. Tc-MYBPA is an Arabidopsis TT2-like transcription factor and functions in the regulation of proanthocyanidin synthesis in *Theobroma cacao*. BMC Plant Biology **15**: 160.

Lizarraga D, Lozano C, Briedé JJ, van Delft JH, Touriño S, Centelles JJ, Torres JL, Cascante M. 2007. The importance of polymerization and galloylation for the antiproliferative properties of procyanidin-rich natural extracts. FEBS Journal **274**: 4802–4811.

Loizzo MR, Pacetti D, Lucci P, Núñez O, Menichini F, Frega NG, Tundis R. 2015. *Prunus persica* var. *platycarpa* (Tabacchiera Peach): Bioactive compounds and antioxidant activity of pulp, peel and seed ethanolic extracts. Plant Foods for Human Nutrition **70**: 331–7.

Lokvam J, Clausen TP, Grapov D, Coley PD, Kursar TA. 2007. Galloyl depsides of tyrosine from young leaves of *Inga laurina*. Journal of Natural Products **70**:134–6.

Lund ST, Bohlmann J. 2006. The molecular basis for wine grape quality--a volatile subject.Science **311**: 804–5.

Luo XD, Basile MJ, Kennelly EJ. 2002. Polyphenolic antioxidants from the fruits of *Chrysophyllum cainito* L. (Star Apple). Journal of Agricultural and Food Chemistry **50**:1379-82.

Lydon J, Duke SO. 1988. Glyphosate induction of elevated levels of hydroxybenzoic acids in higher plants. Journal of Agricultural and Food Chemistry **36**: 813–818.

Ma J, Luo XD, Protiva P, Yang H, Ma C, Basile MJ, Weinstein IB, Kennelly EJ. 2003. Bioactive novel polyphenols from the fruit of *Manilkara zapota* (Sapodilla). Journal of Natural Products **66**: 983–6.

Määttä-Riihinen KR, Kamal-Eldin A, Torronen AR. 2004. Identification and quantification of phenolic compounds in berries of Fragaria and Rubus species (family Rosaceae). Journal of Agricultural and Food Chemistry **52**: 6178–6187.

Macheroux P, Schmid J, Amrhein N, Schaller A. 1999. A unique reaction in a common pathway: mechanism and function of chorismate synthase in the shikimate pathway. Planta **207**: 325–34.

Maeda H, Shasany AK, Schnepp J, Orlova I, Taguchi G, Cooper BR, Rhodes D, Pichersky E, Dudareva N. 2010. RNAi suppression of *Arogenate Dehydratase1* reveals that phenylalanine is synthesized predominantly via the arogenate pathway in petunia petals. Plant Cell **22**: 832-49.

Maeda H, Yoo H, Dudareva N. 2011. Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate. Nature Chemical Biology **7**: 19–21.

Maeda H, Dudareva N. 2012. The shikimate pathway and aromatic amino acid biosynthesis in plants. Annual Review of Plant Biology **63**: 73–105.

Malan JCS, Young DA, Steenkamp JA, Ferreira D. 1988. Oligomeric flavonoids. Part 2. The first profisetinidins with dihydroflavonol consitituent units. Journal of the Chemical Society, Perkin Transactions 1 **9**: 2567-2572.

Mané C, Souquet JM, Ollé D, Verriès C, Véran F, Mazerolles G, Cheynier V, Fulcrand H. 2007. Optimization of simultaneous flavanol, phenolic acid and anthocyanin extraction from grapes using an experimental design; application to the characterization of Champagne grape varieties. Journal of Agricultural Food Chemistry **55**: 7224–7233.

Manpong P, Douglas S, Douglas PL, Pongamphai S, Teppaitoon W. 2009. Preliminary investigation of gallic acid extraction from *Jatropha curcas* Linn. leaves using supercritical carbon dioxide with methanol co-solvent. Journal of Food Process Engineering **34**:1408–1418.

Markham KR, Gould KS, Winefield CS, Mitchella KA, Bloora SJ, Boasec MR. 2000. Anthocyanic vacuolar inclusions-their nature and significance in flower coloration. Phytochemistry **55**: 327–336.

Marinova K, Pourcel L, Weder B, Schwarz M, Barron D, Routaboul JM, Debeaujon I, Klein M. 2007. The Arabidopsis MATE transporter TT12 acts as a vacuolar flavonoid/H+- antiporter active in proanthocyanidin-accumulating cells of the seed coat. Plant Cell **19**: 2023–2038.

Marrs KA, Alfenito MR, Lloyd AM, Walbot V. 1995. A Glutathione-*S*-transferase involved in vacuolar transfer encoded by the maize gene Bronze-2. Nature **375**: 397–400

Masuda T, Iritani K, Yonemori S, Oyama Y, Takeda Y. 2001. Isolation and antioxidant activity of galloyl flavonol glycosides from the seashore plant, *Pemphis acidula*. Bioscience, Biotechnology, and Biochemistry **65**:1302-9.

Matthews S, Mila I, Scalbert A, Donnelly DMX. 1997. Extractable and non-extractable proanthocyanidins in barks. Phytochemistry **45**: 405-410.

Matthews MA, Shackel KA. 2005. Growth and water transport in fleshy fruit. In: Holbrook NM, Zwieniecki MA, eds. Vascular transport in plants. Burlington, California: Academic Press, 189–197.

Mattivi F, Guzzon R, Vrhovsek U, Stefanini M, Velasco R. 2006. Metabolite profiling of grape: Flavonols and anthocyanins. Journal of Agricultural and Food Chemistry **54**: 7692–7702.

Matus JT, Aquea F, Arce-Johnson P. 2008. Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and Arabidopsis genomes. BMC Plant Biology **8**: 83.

May P. 2000. From bud to berry, with special reference to inflorescence and bunch morphology in *Vitisvinifera* L. Australian Journal of Grape and Wine Research **6**: 82–98.

Mazza G, Fukumoto L, Delaquis P, Girard B, Ewert B. 1999. Anthocyanins, phenolics, and color of Cabernet Franc, Merlot, and Pinot Noir wines from British Columbia. Journal of Agricultural Food Chemistry **47**:4009-17.

McCarthy MG, Coombe BG. 1999. Is weight loss in ripening grape berries cv. Shiraz caused by impeded phloem transport? Australian Journal of Grape and Wine Research **5**: 17–21.

McDonald MS, Hughes M, Burns J, Lean ME, Matthews D, Crozier A. 1998. Survey of the free and conjugated myricetin and quercetin content of red wines of different geographical origins. Journal of Agricultural Food Chemistry **46**: 368-375.

McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R. 2013. Analysis Tool Web Services from the EMBL-EBI. Nucleic acids research 41 (Web Server issue):W597-600.

Mehrtens F, Kranz H, Bednarek P, Weisshaar B. 2005. The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. Plant Physiology **138**: 1083-96.

Mellway RD, Tran LT, Prouse MB, Campbell MM, Constabel CP. 2009. The wound-, pathogen-, and ultraviolet B-responsive MYB134 gene encodes an R2R3 MYB transcription factor that regulates proanthocyanidin synthesis in poplar. Plant Physiology **150**: 924–941.

Menting JGT, Scopes RK, Stevenson TW. 1994. Characterization of flavonoid 3',5'-hydroxylase in microsomal membrane-fraction of *Petunia hybrida* flowers. Plant Physiology **106**: 633–642.

Mertz C, Cheynier V, Günata Z, Brat P. 2007. Analysis of phenolic compounds in two blackberry species (*Rubus glaucus* and *Rubus adenotrichus*) by high-performance liquid chromatography with diode array detection and electrospray ion trap mass spectrometry. Journal of Agricultural Food Chemistry. **55**: 8616–24.

Michel G, Roszak AW, Sauvé V, Maclean J, Matte A, Coggins JR, Cygler M, Lapthorn AJ. 2003. Structures of shikimate dehydrogenase AroE and its Paralog YdiB. A common structural framework for different activities. The Journal of Biological Chemistry **278**:19463-72.

Milkowski C, Baumert A, Schmidt D, Nehlin L, Strack D. 2004. Molecular regulation of sinapate ester metabolism in *Brassica napus*: Expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. The Plant Journal **38**: 80–92.

Milo R, Last RL. 2012. Achieving diversity in the face of constraints: lessons from metabolism. Science **336**: 1663-1667.

Min KJ, Kwon TK. 2014. Anticancer effects and molecular mechanisms of epigallocatechin-3-gallate. Integrative Medicine Research **3**: 16–24.

Miranda M, Ralph SG, Mellway R, White R, Heath MC, Bohlmann J, Constabel CP. 2007. The transcriptional response of hybrid poplar (Populus trichocarpa x P. deltoides) to infection by Melampsora medusae leaf rust involves induction of flavonoid pathway genes leading to the accumulation of proanthocyanidins. Mol Plant Microbe Interactions **20**: 816-31.

Mizutani M, Ohta D, Sato R. 1997. Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from Arabidopsis and its expression manner in planta. Plant Physiology **113**: 755-63.

Monagas M, Gomez-Cordoves C, Bartolomé B, Laureano O, Ricardo da Silva J. 2003. Monomeric, oligomeric, and polymeric flavan-3-ol composition of wines and grapes from *Vitis vinifera* L. Cv Graciano, Tempranillo, and Cabernet Sauvignon. Journal of Agricultural Food Chemistry **51**: 6475–6481.

Monteiro F, Sebastiana M, Pais MS, Figueiredo A. 2013. Reference gene selection and validation for the early responses to downy mildew infection in susceptible and resistant *Vitis vinifera* cultivars. PLoS One **8**: e72998.

Moore Bd. 2004. Bifunctional and moonlighting enzymes: lighting the way to regulatory control. Trends in Plant Science **9**:221-8.

Moore JP, Westall KL, Ravenscroft N, Farrant JM, Lindsey GG, Brandt WF. 2005. The predominant polyphenol in the leaves of the resurrection plant *Myrothamnus flabellifolius*, 3,4,5 tri-*O*-galloylquinic acid, protects membranes against desiccation and free radical-induced oxidation. Biochemistry Journal **385**: 301–8.

Mori K, Sugaya S, Gemma H. 2005. Decreased anthocyanin biosynthesis in grape berries grown under elevated night temperature condition. Scientia Horticulturale **105**: 319–330.

Moriyama Y, Hiasa M, Matsumoto T, Omote H. 2008. Multidrug and toxic compound extrusion (MATE)-type proteins as anchor transporters for the excretion of metabolic waste products and xenobiotics. Xenobiotica **38**: 1107–18.

Moura DS, Bergey DR, Ryan CA. 2001. Characterization and localization of a wound-inducible type I serine-carboxypeptidase from leaves of tomato plants (*Lycopersicon esculentum* Mill.). Planta **212**: 222-30.

Mousdale DM, Coggins JR. 1985. Subcellular localization of the common shikimate-pathway enzymes in *Pisum sativum* L. Planta **163**: 241–249.

Mousdale DM, Campbell MS, Coggins JR. 1987. Purification and characterization of a bifunctional dehydroquinase-shikimate: NADP oxidoreductase from pea seedlings. Phytochemistry **26**: 2665–2670.

Moustafa E, Wong E. 1967. Purification and properties of chalcone-flavanone isomerase from soybean seed. Phytochemistry **6**: 625–632.

Mueller LA, Walbot V. 2001. Models for vacuolar sequestration of anthocyanins. Regulation of Phytochemicals by Molecular Techniques **35**: 297-312.

Mugford ST, Qi X, Bakht S, Hill L, Wegel E, Hughes RK, Papadopoulou K, Melton R, Philo M, Sainsbury F, Lomonossoff GP, Roy AD, Goss RJ and Osbourn A. 2009. A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. Plant Cell **21:** 2473–2484.

Mugford ST, Louveau T, Melton R, Qi X, Bakht S, Hill L,Tsurushima T, Honkanen S, Rosser SJ, Lomonossoff GP, Osbourn A. 2013. Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat.The Plant Cell **25**: 1078–1092.

Muir RM, Ibáñez AM, Uratsu SL, Ingham ES, Leslie CA, McGranahan GH, Batra N, Goyal S, Joseph J, Jemmis ED, Dandekar AM. 2011. Mechanism of gallic acid biosynthesis in bacteria (*Escherichia coli*) and walnut (*Juglans regia*). Plant Molecular Biology **75**:555–65.

Nair RB, Bastress KL, Ruegger MO, Denault JW, Chapple C. 2004. The *Arabidopsis thaliana* REDUCED EPIDERMAL FLUORESCENCE1 gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis. Plant Cell **16**: 544–54.

Ncube EN, Mhlongo MI, Piater LA, Steenkamp PA, Dubery IA, Madala NE. 2014. Analyses of chlorogenic acids and related cinnamic acid derivatives from *Nicotiana tabacum* tissues with the aid of UPLC-QTOF-MS/MS based on the in-source collision-induced dissociation method. Chemical Central Journal **8**: 66.

Nesi N, Jond C, Debeaujon I, Caboche M, Lepiniec L. 2001. The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. Plant Cell **13**: 2099-114.

Niemetz R, Gross GG. 2005. Enzymology of gallotannin andellagitannin biosynthesis. Phytochemistry **66**: 2001–2011.

Niggeweg R, Michael AJ, Martin C. 2004. Engineering plants with increased levels of the antioxidant chlorogenic acid. Nature Biotechnology **22**: 746–754.

Nimal Punyasiri PA, Tanner GJ, Abeysinghe IS, Kumar V, Campbell PM, Pradeepa NH. 2004. *Exobasidium vexans* infection of *Camellia sinensis* increased 2,3-cis isomerisation and gallate esterification of proanthocyanidins. Phytochemistry **65**: 2987-94.

Nishizaki Y, Yasunaga M, Okamoto E, Okamoto M, Hirose Y, Yamaguchi M, Ozeki Y, Sasaki N. 2013. p-Hydroxybenzoyl-glucose is a zwitter donor for the biosynthesis of 7-polyacylated anthocyanin in Delphinium. Plant Cell **25**:4150–4165.

Noda N, Kazuma K, Sasaki T, Furukawa K, Suzumi M. (Aomori prefecture) novel aromatic acyl group transferase gene. Patent Publication Number WO/2007/046148, April 26, 2007.

Nützmann HW, Osbourn A. 2014. Gene clustering in plant specialized metabolism. Current Opinion in Biotechnology **26**: 91-9.

Ohnishi M, Hirose S, Kawaguchi M, Ito S, Fujino Y. 1990. Chemical composition of lipids, especially triacylglycerol, in grape seeds.Agricultural and Biological Chemistry **54**:1035-1042.

Ojeda H, Deloire A, Carbonneau A, Ageorges A, Romieu C. 1999. Berry development of grapevines: Relations between the growth of berries and their DNA content indicate cell multiplication and enlargement. Vitis **38**: 145–150.

Ojeda H, Andary C, Kraeva E, Carbonneau A, Deloire A. 2002. Influence of pre- and postvéraison water deficit on synthesis and concentration of skin phenolic compounds during berry growth of *Vitis vinifera* cv. Shiraz. American Journal of Enology and Viticulture **53**: 261-267.

Ollat N, Diakou-Verdin P, Carde JP, Barrieu F, Gaudillère JP, Moing A. 2002. Grapeberrydevelopment: areview. Journal International des Sciences de la Vigne et du Vin **36**: 109–131.

Ollé D, Guiraud JL, Souquet JM, Terrier N, Ageorges A, Cheynier V, Verries C. 2011. Effect of pre- and post-veraison water deficit on proanthocyanidin and anthocyanin accumulation during Shiraz berry development. Australian Journal of Grape and Wine Research **17**: 90–100.

Ohnishi M, Hirose S, Kawaguchi M, Ito S, Fujino Y. 1990. Chemical composition of lipids, especially triacylglycerol, in grape seeds. Agricultural and Biological Chemistry **54**: 1035-1042.

Ono E, Homma Y, Horikawa M, Kunikane-Doi S, Imai H, Takahashi S, Kawai Y, Ishiguro M, Fukui Y, Nakayama T. 2010. Functional differentiation of the glycosyltransferases that contribute to the chemical diversity of bioactive flavonol glycosides in grapevines (*Vitis vinifera*). Plant Cell **22**: 2856-71.

Onyeneho SN, Hettiarachchy NS. 1993. Antioxidant activity, fatty acids and phenolic acids compositions of potato peels. Journal of the Science of Food and Agriculture **62**: 345–350.

Osawa Y. 1982. Copigmentation of anthocyanins. In Anthocyanins as Food Colours, ed. P. Markakis, Academic Press: New York, pp. 41-68.

Olschläger C, Regos I, Zeller FJ, Treutter D. 2008. Identification of galloylated propelargonidins and procyanidins in buckwheat grain and quantification of rutin and flavanols from homostylous hybrids originating from *F. esculentum* x *F. homotropicum*. Phytochemistry **69**:1389–97.

Osbourn A. 2010. Gene clusters for secondary metabolic pathways: an emerging theme in plant biology. Plant Physiology **154:** 531-5.

Osman H, Nasarudin R, Lee SL. 2004. Extracts of cocoa (*Theobroma cacao* L.) leaves and their antioxidation potential. Food Chemistry **86**: 41–46.

Ossipov V, Salminen JP, Ossipova S, Haukioja E, Pihlaja K. 2003. Gallic acid and hydrolysable tannins are formed in birch leaves from an intermediate compound of the shikimate pathway. Biochemical Systematics and Ecology **31**: 3–16.

Palafox-Carlos H, Yahia EM, González-Aguilar GA. 2012. Identification and quantification of major phenolic compounds from mango (*Mangifera indica*, cv. Ataulfo) fruit by HPLC–DAD–MS/MS-ESI and their individual contribution to the antioxidant activity during ripening. Food Chemistry **135**: 105–111.

Pang Y, Peel GJ, Wright E, Wang Z, Dixon RA. 2007. Early steps in proanthocyanidin biosynthesis in the model legume *Medicago truncatula*. Plant Physiology **145**: 601-615.

Pang Y, Peel GJ, Sharma SB, Tang Y, Dixon RA. 2008. A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of *Medicago truncatula*. Proceedings of the National Academy of Sciences of the United States of America **105**:14210–14205.

Pang Y, Cheng X, Huhman DV, Ma J, Peel GJ, Yonekura-Sakakibara K, Saito K, Shen G, Sumner LW, Tang Y, Wen J, Yun J, Dixon RA. 2013. Medicago glucosyltransferase UGT72L1: potential rolesin proanthocyanidin biosynthesis. Planta **238**:139-154.

Park CH, Tanaka T, Kim HY, Park JC, Yokozawa T. 2012. Protective effects of corni fructus against advanced glycation endproducts and radical scavenging. Evidence-Based Complementary Alternative Medecine **2012**:418953.

Passamonti S, Cocolo A, Braidot E, Petrussa E, Peresson C, Medic N, Macri F, Vianello A. 2005. Characterization of electrogenic bromosulphophthalein transport in carnation petal microsomes and its inhibition by antibodies against bilitranslocase. FEBS Journal **272**: 3282-3296.

Pastrana-Bonilla E, Akoh CC, Sellappan S, Krewer G. 2003. Phenolic content and antioxidant capacity of muscadine grapes. Journal of Agricultural and Food Chemistry **51**:5497–5503.

Pathak SB, Niranjan K, Padh H, Rajani M. 2004. TLC densitometric method for the quantification of eugenol and gallic acid in clove. Chromatographia **60**:241-244.

Peer WA, Murphy AS. 2007. Flavonoids and auxin transport: modulators or regulators? Trends in Plant Science **12**: 556–563.

Pereira GE, Gaudillere JP, Pieri P, Hilbert G, Maucourt M, Deborde C, Moing A, Rolin D. 2006. Microclimate influence on mineral and metabolic profiles of grape berries. Journal of Agricultural and Food Chemistry **54**: 6765-75.

Pérez-Díaz R, Ryngajllo M, Pérez-Díaz J, Peña-Cortés H, Casaretto JA, González-Villanueva E, Ruiz-Lara S. 2014. VvMATE1 and VvMATE2 encode putative proanthocyanidin transporters expressed during berry development in *Vitis vinifera* L. Plant Cell Reports **33**: 1147-59.

Peters DJ, Constabel CP. 2002. Molecular analysis of herbivore-induced condensed tannin synthesis: cloning and expression of dihydroflavonol reductase from trembling aspen (*Populus tremuloides*). The Plant Journal **32**: 701–712.

Petit P, Granier T, d'Estaintot BL, Manigand C, Bathany K, Schmitter JM, Lauvergeat V, Hamdi S, Gallois B. 2007. Crystal structure of grape dihydroflavonol 4-reductase, a key enzyme in flavonoid biosynthesis. Journal of Molecular Biology **368**: 1345-57.

Petit AN, Baillieul F, Vaillant-Gaveau N, Jacquens L, Conreux A, Jeandet P, Clément C, Fontaine F. 2009. Low responsiveness of grapevine flowers and berries at fruit set to UV-C irradiation. Journal of Experimental Biology **60**: 1155-1162.

Pezet R, Perret C, Jean-Denis JB, Tabacchi R, Gindro K, Viret O. 2003. δ-viniferin, a resveratrol dehydrodimer: one of the major stilbenes synthesized by grapevine leaves. Journal of Agricultural and Food Chemistry **51**:5488–5492.

Pezet R, Gindro K, Viret O, Richter H. 2004. Effects of resveratrol, viniferins and pterostilbene on *Plasmopara viticola* zoospore mobility and disease development. Vitis **43**: 145–148.

Pinelo M, Arnous A, Meyer AS. 2006. Upgrading of grape skins: Significance of plant cell-wall structural components and extraction techniques for phenol release. Trends in Food Science & Technology **17**: 579–590.

Ping L, Xu-Qing W, Huai-Zhou W, Yong-Ning W. 1993. High performance liquid chromatographic determination of phenolic acids in fruits and vegetables. Biomedical and Environmental Sciences **6**: 389-398.

Pourcel L, Routaboul JM, Kerhoas L, Caboche M, Lepiniec L, Debeaujon I. 2005. TRANSPARENT TESTA10 encodes a laccase-like enzyme involved in oxidative polymerization of flavonoids in Arabidopsis seed coat. Plant Cell **17**: 2966–2980.

Pourcel L, Irani NG, Lu Y, Riedl K, Schwartz S, Grotewold E. 2010. The formation of Anthocyanic Vacuolar Inclusions in *Arabidopsis thaliana* and implications for the sequestration of anthocyanin pigments. Molecular Plant **3:** 78-90.

Poustka F, Irani NG, Feller A, Lu Y, Pourcel L, Frame K, Grotewold E. 2007. A trafficking pathway for anthocyanins overlaps with the endoplasmic reticulum-to-vacuole protein sorting route in Arabidopsis and contributes to the formation of vacuolar inclusions. Plant Physiology **145**: 1323-1335.

Price SF, Breen PJ, Valladao M, Watson BT. 1995. Cluster sun exposure and quercetin in Pinot Noir grapes and wine.American Journal of Enology and Viticulture **46**: 187-194.

Prieur C, Rigaud J, Cheynier V, Moutounet M. 1994. Oligomeric and polymeric procyanidins from grape seeds. Phytochemistry **36**: 781–784.

Puech JL, Mertz C, Michon V, Le Guernevé C, Doco T, Hervé du Penhoat C. 1999. Evolution of castalagin and vescalagin in ethanol solutions. Identification of new derivatives. Journal of Agricultural and Food Chemistry **47**: 2060–2066.

Pyo Y, Lee TC, Logendra L, Rosen RT. 2004. Antioxidant activity and phenolic compounds of Swiss chard (*Beta vulgaris* subspecies *cycla*) extracts. Food Chemistry **85**: 19–26.

Pyrzynska K, Biesaga M. 2009. Analysis of phenolic acids and flavonoids in honey.TrAC Trends in Analytical Chemistry **28**: 893–902.

Qu W, BreksaIii AP, Pan Z, Ma H. 2012. Quantitative determination of major polyphenol constituentsin pomegranate products. Food Chemistry **132**: 1585–1591.

Quideau S, Deffieux D, Douat-Casassus C, Pouységu L. 2011. Plant polyphenols: chemical properties, biological activities, and synthesis. Angewandte Chemie (International Edition in English) **50**: 586–621.

Rani PU, Pratyusha S. 2013. Defensive role of *Gossypium hirsutum* L. anti-oxidative enzymes and phenolic acids in response to *Spodoptera litura* F. feeding. Journal of Asia-Pacific Entomology **16**: 131–136.

Ravaglia D, Espley RV, Henry-Kirk RA, Andreotti C, Ziosi V, Hellens RP, Costa G, Allan AC. 2013. Transcriptional regulation of flavonoid biosynthesis in nectarine (*Prunus persica*) by a set of R2R3 MYB transcription factors. BMC Plant Biology **13:** 68.

Rech SS, Heidt S, Requena N. 2013. A tandem Kunitz protease inhibitor (KPI106)-serine carboxypeptidase (SCP1) controls mycorrhiza establishment and arbuscule development in Medicago truncatula. Plant Journal **75**: 711-25.

Reddivari L, Hale AL, Miller JC Jr. 2007. Determination of phenolic content, composition and their contribution to antioxidant activity in specialty potato selections. American Journal of Potato Research **84**:275-282.

Reiersen B, Kiremire BT, Byamukama R, Andersen ØM. 2003. Anthocyanins acylated with gallic acid from chenille plant, *Acalypha hispida*. Phytochemistry **64**: 867–871.

Rentzsch M, Wilkens A, Winterhalter P. 2009. Non-flavonoid phenolic compounds. In: Moreno-Arribas, M.V., Polo, M.C. (Eds.), Wine Chemistry and Biochemistry. Springer, USA, pp. 509–528.

Ribéreau-Gayon P. 1964. Les composés phénoliques du raisin et du vin, Institut National de la Recherche Agronomique, Paris.

Richman JE, Chang YC,Kambourakis S, Draths KM,Almy E, Snell KD, Strasburg GM, Frost JW. 1996. Reaction of 3-dehydroshikimic acid with molecular oxygen and hydrogen peroxide: products, mechanism, and associated antioxidant activity. Journal of American Chemical Society **118**:11587-11591.

Rinaldo A, Cavallini E, Jia Y, Moss SM, McDavid DA, Hooper LC, Robinson SP, Tornielli GB, Zenoni S, Ford CM, Boss PK, Walker AR. 2015. A grapevine anthocyanin acyltransferase, transcriptionally regulated by VvMYBA, can produce most acylated anthocyanins present in grape skins. Plant Physiology pii: pp.01255.2015.

Rizzon C, Ponger L, Gaut BS. 2006. Striking similarities in the genomic distribution of tandemly arrayed genes in Arabidopsis and rice. PLoS Comput Biology **2**:e115.

Roggero JP, Coen S, Ragonnet B. 1986. High performance liquid chromatography survey on changes in pigment content in ripening grapes of Syrah. An approach to anthocyanin metabolism. American Journal of Enological Viticulture **37**: 77-83.

Roggero JP, Garcia-Parrilla C. 1995. Effects of ultraviolet irradiation on resveratrol and changes in resveratrol and various of its derivatives in the skins of ripening grapes. Science des Aliments **15**:411–422.

Romani A, Pinelli P,Galardi C, Mulinacci N, Tattini M. 2002. Identification and quantification of galloyl derivatives, flavonoid glycosides and anthocyanins in leaves of *Pistacia lentiscus* L. Phytochemical Analysis **13**: 79–86.

Romeyer FM, Macheix JJ, Goiffon JJ, Reminac CC, Sapis JC. 1983. The browning capacity ofgrapes. 3. Changes and importance of hydroxycinnamic acid tartaric acid esters duringdevelopment and maturation of the fruit. Journal of Agriculture and Food Chemistry **31**: 346-349.

Rossetto M, Lante A, Vanzani P, Spettoli P, Scarpa M, Rigo A. 2005. Red chicories as potent scavengers of highly reactive radicals: a study on their phenolic composition and peroxyl radical trapping capacity and efficiency. Journal of Agricultural and Food Chemistry **53**: 8169-8175.

Routaboul JM, Kerhoas L, Debeaujon I, Pourcel L, Caboche M, Einhorn J, Lepiniec L. 2006. Flavonoid diversity and biosynthesis in seed of *Arabidopsis thaliana*. Planta **224**:96-107.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology **132**: 365-86.

Rudrappa T, Bonsall J, Gallagher JL, Seliskar DM, Bais HP. 2007. Root-secreted allelochemical in the noxious weed *Phragmites Australis* deploys a reactive oxygen species response and microtubule assembly disruption to execute rhizotoxicity. Journal of Chemical Ecology **33**:1898–1918.

Ruffner HP. 1982. Metabolism of tartaric and malic acids in Vitis: a review-part b. Vitis **21**: 346-358.

Russell DW, Conn EE. 1967. The cinnamic acid 4-hydroxylase of pea seedlings: isolation, activity, cofactor requirements. Archive of Biochemistry and Biophysics **122**: 256–258.

Saijo R, Nonaka GI, Nishioka I. 1990. Gallic acid esters of bergenin and norbergenin from *Mallotus japonicus*. Phytochemistry **29**: 267–270.

Sainsbury F, Thuenemann EC, Lomonossoff GP. 2009. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnol Journal **7**: 682–93.

Saito K. 2013. Phytochemical genomics--a new trend. Current Opinion in Plant Biology **16**: 373-80.

Saito N, Tatsuzawa F, Miyoshi K, Shigihara A, Honda T. 2003. The first isolation of C-glycosylanthocyanin from the flowers of *Tricyrtis formosana*. Tetrahedron Letters **44**: 6821-6823.

Salas E, Dueñas M, Schwarz M, Winterhalter P, Cheynier V, Fulcrand H. 2005. Characterization of pigments from different high speed countercurrent chromatography wine fractions. Journal of Agricultural Food Chemistry **53**: 4536–4546.

Saldanha LL, Vilegas W, Dokkedal AL. 2013. Characterization of flavonoids and phenolic acids in *Myrcia bella* Cambess. using FIA-ESI-IT-MS(n) and HPLC-PAD-ESI-IT-MS combined with NMR. Molecules.**18**: 8402–16.

Santos SA, Freire CS, Domingues MR, Silvestre AJ, Pascoal Neto C. 2011. Characterization of phenolic components in polar extracts of *Eucalyptus globulus* Labill. bark by high-performance liquid chromatography-mass spectrometry. Journal of Agricultural Food Chemistry **59**: 9386–93.

Santos-Rosa M, Poutaraud A, Merdinoglu D, Mestre P. 2008. Development of a transient expression system in grapevine via agro-infiltration. Plant Cell Reports **27**: 1053–63.

Sanz M, Cadahía E, Esteruelas E, Muñoz AM, Fernández De Simón B, Hernández T, Estrella I. 2010. Phenolic compounds in cherry (*Prunus avium*) heartwood with a view to their use in cooperage. Journal of Agricultural Food Chemistry **58**: 4907–14.

Sanz M, Fernández de Simón B, Esteruelas E, Muñoz ÁM, Cadahía E, Hernández T, Estrella I, Pinto E. 2011. Effect of toasting intensity at cooperage on phenolic compounds in acacia (*Robinia pseudoacacia*) heartwood. Journal of Agricultural Food Chemistry **59**: 3135–45.

Schaart J G, Dubos C, Romero De La Fuente I, van Houwelingen A M, de Vos R C, Jonker H H, Xu W, Routaboul J M, Lepiniec L, Bovy AG. 2013. Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria* x *ananassa*) fruits. New Phytologist **197**: 454–467.

Schmid J, Schaller A, Leibinger U, Boll W, Amrhein N. 1992. The in vitro synthesized tomato shikimate kinase precursor is enzymatically active and is imported and processed to the matureenzyme by chloroplasts. The Plant Journal**2**: 375–383.

Schmid J, Amrhein N. 1999. The shikimate pathway. In: Singh BK (ed) Plant Amino Acids, pp 147–169. Marcel Dekker, New York.

Schmitzer V, Slatnar A, Veberic R, Stampar F, Solar A. 2011. Roasting affects phenolic composition and antioxidative activity of hazelnuts (*Corylus avellana* L.). Journal of Food Science **76**: S14-9.

Schwarz M, Winterhalter P. 2003. A novel synthetic route to substituted pyranoanthocyanins with unique colour properties. Tetrahedron Letters **44**: 7583–7587.

Serrano M, Wang B, Aryal B, Garcion C, Abou-Mansour E, Heck S,Geisler M, Mauch F, Nawrath C, Métraux JP. 2013. Export of salicylic acid from the chloroplast requires the multidrug and toxin extrusion-like transporter EDS5. Plant Physiology **162**: 1815–1821.

Sgarbi E, Fornasiero RB, Lins AP, Bonatti PM. 2003. Phenol metabolism is differentially affected by ozone in two cell lines from grape (*Vitis vinifera* L.) leaf. Plant Science **165**: 951–957.

Shan B, Cai YZ, Sun M, Corke H. 2005. Antioxidant capacity of 26 spice extracts and characterization of their phenolic constituents. Journal of Agricultural Food Chemistry **53**: 7749–59.

Shirley AM, Chapple C. 2003. Biochemical characterization of sinapoyl:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. Journal of Biological Chemistry **278**: 19870–19877.

Simmonds MS. 2003. Flavonoid-insect interactions: recent advances in our knowledge. Phytochemistry **64**:21-30.

Singh SA, Christendat D. 2006. Structure of Arabidopsis dehydroquinate dehydratase-shikimate dehydrogenase and implications for metabolic channeling in the shikimate pathway. Biochemistry **45**: 7787-96.

Singh SA, Christendat D. 2007. The DHQ-dehydroshikimate-SDH-shikimate-NADP(H) complex: insights into metabolite transfer in the shikimate pathway.Crystal Growth & Design **7**:2153–2160.

Singh S, Stavrinides J, Christendat D, Guttman DS. 2008. A phylogenomic analysis of the shikimate dehydrogenases reveals broadscale functional diversification and identifies one functionally distinct subclass. Molecular Biology and Evolution **25**:2221–2232.

Singh H, Dixit S, Verma PC, Singh PK. 2014. Evaluation of total phenolic compounds and insecticidal and antioxidant activities of tomato hairy root extract. Journal of Agricultural Food Chemistry **62**:2588-94.

Singleton VL, Timberlake CF, Lea AGH. 1978. The phenolic cinnamates of grapes and wine. Journal of Sciences and Food Agriculture **29**: 403-410.

Smith GJ, Markham KR. 1998. Tautomerism of flavonol glucosides: relevance to plant UV protection and flower colour. Journal of Photochemistry and Photobiology A: Chemistry **118**: 99–105.

Song CP, Guo Y, Qiu Q, Lambert G, Galbraith DW, Jagendorf A, Zhu JK. 2004. A probable $Na^+(K^+)/H^+$ exchanger on the chloroplast envelope functions in pH homeostasis and chloroplast development in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences **101**:10211–6.

Sørensen SB, Breddam K, Svendsen I. 1986. Primary structure of carboxypeptidase I from malted barley. Carlsberg Research Communications **51**: 475-485.

Sørensen SB, Svendsen I, Breddam K.1987. Primary structure of carboxypeptidase II from malted barley. Carlsberg Research Communications **52**: 285-295.

Sørensen SB, Svendsen I, Breddam K. 1989. Primary structure of carboxypeptidase III from malted barley. Carlsberg Research Communications **54**: 193-202.

Sørensen SB, Breddam K. 1997. The specificity of carboxypeptidase Y may be altered by changing the hydrophobicity of the S'1 binding pocket. Protein Science **6**: 2227–2232.

Souquet JM, Cheynier V, Brossaud F, Moutounet M. 1996. Polymeric proanthocyanidins from grape skins. Phytochemistry **43**: 509-512.

Souquet JM, Labarbe B, Le Guernevé C, Cheynier V, Moutounet, M. 2000. Phenolic composition of grape stems. Journal of Agricultural Food Chemistry **48**: 1076-1080.

Souquet JM, Veran F, Mané C, Cheynier V. 2006. Optimization of extraction conditions on phenolic yields from the different parts of grape clusters – quantitative distribution of their proanthocyanidins. Proceedings of the XXIII International Conference on Polyphenols,Winnipeg, Manitoba, Canada.

Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C. 1994. Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.) Plant Molecular Biology, **24**: 743–755.

St Pierre B, Laflamme P, Alarco AM, De Luca V. 1998. The terminal *O*-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A dependent acyl transfer. *The Plant Journal* **14**: 703–713.

Stafford HA, Lester HH. 1982. Enzymatic and non enzymatic reduction of (+)-dihydroquercetin to its 3,4,-diol. Plant Physiology **70**: 695–698.

Stehle F, Brandt W, Milkowski C, Strack D. 2006. Structure determinants and substrate recognition of serine carboxypeptidase-like acyltransferases from plant secondary metabolism. FEBS Letters **580**: 6366–74.

Stehle F, Stubbs MT, Strack D, Milkowski C. 2008. Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism. The FEBS Journal **275**: 775–787.

Stehle F, Brandt W, Stubbs MT, Milkowski C, Strack D. 2009. Sinapoyltransferases in the light of molecular evolution. Phytochemistry **70**: 1652-62.

Strack D, Wray V, Metzger JW, Grosse W. 1992. Two anthocyanins acylated with gallic acid from the leaves of *Victoria amazonica*. Phytochemistry **31**: 989-991.

Stracke R, Werber M, Weisshaar B. 2001. The R2R3-MYB gene family in Arabidopsis thaliana. Current Opinion in Plant Biology **4**: 447-56.

Su CT, Singleton V. 1969. Identification of three flavan-3-ols from grapes. Phytochemistry **8**: 1553–1558.

Sullivan ML. 2009. A novel red clover hydroxycinnamoyl transferase has enzymatic activities consistent with a role in phaselic acid biosynthesis. Plant Physiology **150**: 1866–1879.

Sullivan ML, Zeller WE. 2012. Efficacy of various naturally occurring caffeic acid derivatives in preventing post-harvest protein losses in forages. Journal of the Science of Food and Agriculture **93**: 219–226.

Sullivan ML. 2014. Perennial peanut (*Arachis glabrata* Benth.) leaves contain hydroxycinnamoyl-CoA:tartaric acid hydroxycinnamoyl transferase activity and accumulate hydroxycinnamoyl-tartaric acid esters. Planta **239**: 1091-100.

Suzich JA, Dean JFD, Herrmann KM. 1985. 3-Deoxy-D-arabino-Heptulosonate 7-Phosphate Synthase from carrot root (*Daucus carota*) is a hysteretic enzyme. Plant Physiology **79**: 765–770.

Symons GM, Davies C, Shavrukov Y, Dry IB, Reid JB, Thomas MR. 2006. Grapes on steroids. Brassinosteroids are involved in grape berry ripening. Plant Physiology **140**: 150-158.

Tako E, Beebe SE, Reed S, Hart JJ, Glahn RP. 2014. Polyphenolic compounds appear to limit the nutritional benefit of biofortified higher iron black bean (*Phaseolus vulgaris* L.). Nutrition Journal **13**: 28.

Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution **30**: 2725-2729.

Taneyama M. 1992. Studies on C-Glycosides in higher plants intracellular distribution of bergenin and related compounds in *Saxifraga stolonifera* leaves.The botanical magazine, Tokyo **105**: 565–571.

Tanner GJ, Francki KT, Abrahams S,Watson JM, Larkin PJ, Ashton AR. 2003. Proanthocyanidin biosynthesis in plants. Purification of legume leucoanthocyanidin reductase and molecular cloning of its cDNA. Journal of Biological Chemistry **278**: 31647–31656.

Taylor LP, Jorgensen R. 1992. Conditional male fertility in chalcone synthase-deficient petunia. Journal of Heredity **83**:11–17.

Terrier N, Sauvage FX, Ageorges A, Romieu C. 2001. Changes in acidity and in proton transport at the tonoplast of grape berries during development. Planta **213**: 20–28.

Terrier N, Glissant D, Grimplet J, Barrieu F, Abbal P, Couture C, Ageorges A, Atanassova R, Léon C, Renaudin JP, Dédaldéchamp F, Romieu C, Delrot S, Hamdi S. 2005. Isogene specific oligo arrays reveal multifaceted changes in gene expression during grape berry (*Vitis vinifera* L.) development. Planta **222**: 832-47.

Terrier N, Torregrosa L, Ageorges A, Vialet S, Verries C, Cheynier V, Romieu C. 2009a. Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. Plant Physiology **149**: 1028–1041.

Terrier N, Ollé D, Verriès C, Cheynier V. 2009b. Biochemical and molecular aspects of flavan-3-ol synthesis during berry development. In Grapevine Molecular Physiology and Biotechnology. Edited by: Roubelakis-Angelakis KA, Roubelakis-Angelakis KA. Netherlands: Springer 365-388.

Tesnière C, Torregrosa L, Pradal M, Souquet JM, Gilles C, Dos Santos K, Chatelet P, Gunata Z. 2006. Effects of genetic manipulation of alcohol dehydrogenase levels on the response to stress and the synthesis of secondary metabolites in grapevine leaves. Journal of Experimental Botany **57**: 91-9.

Thipyapong P, Hunt MD, Steffens JC. 2004. Antisense downregulation of polyphenol oxidase results in enhanced disease susceptibility. Planta **220**:105–117.

Thompson EP, Wilkins C, Demidchik V, Davies JM, Glover BJ. 2010. An Arabidopsis flavonoid transporter is required for anther dehiscence and pollen development. Journal of Experimental Botany **61**: 439–451.

Tian L,Wan SB, Pan QH, Zheng YJ, Huang WD. 2008. A novel plastid localization of chalcone synthase in developping grape berry. Plant Science **175**: 431–436.

Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M, Noji M, Yamazaki M, Saito K. 2005. Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor. Plant Journal **42**: 218-235.

Tohge T, Watanabe M, Hoefgen R, Fernie AR. 2013. Shikimate and phenylalanine biosynthesis in the green lineage. Frontiers in Plant Science **4**: 62.

Tomaino A, Martorana M, Arcoraci T, Monteleone D, Giovinazzo C, Saija A. 2010. Antioxidant activity and phenolic profile of pistachio (*Pistacia vera* L., variety Bronte) seeds and skins. Biochimie **92**:1115-22.

Torregrosa L, Bouquet A. 1997. Agrobacterium rhizogenes and A. tumefaciens co-transformation to obtain grapevine hairy roots producing the coat protein of grapevine chrome mosaic nepovirus. Plant Cell, Tissue and Organ Culture **49**: 53-62.

Touriño S, Lizárraga D, Carreras A, Lorenzo S, Ugartondo V, Mitjans M, Vinardell MP, Juliá L, Cascante M, Torres JL. 2008. Highly galloylated tannin fractions from witch hazel (*Hamamelis virginiana*) bark: electron transfer capacity, in vitro antioxidant activity, and effects on skin-related cells. Chemical Research in Toxicology **21**: 696–704.

TzinV, Galili G. 2010. New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. Molecular Plant **3**: 956–972.

Vallarino JG, Osorio S, Bombarely A, Casañal A, Cruz-Rus E, Sánchez-Sevilla JF, Amaya I, Giavalisco P, Fernie AR, Botella MA, Valpuesta V. 2015. Central role of FaGAMYB in the transition of the strawberry receptacle from development to ripening. New Phytologist **208**: 482-96.

Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W. 2010. Lignin biosynthesis and structure. Plant Physiology **153**: 895-905.

Veberic R, Colaric M, Stampar F. 2008. Phenolic acids and flavonoids of fig fruit (*Ficus carica* L.) in the northern Mediterranean region. Food Chemistry **106**: 153–157.

Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, Fitzgerald LM, Vezzulli S, Reid J, Malacarne G, Iliev D, Coppola G, Wardell B, Micheletti D, Macalma T, Facci M, Mitchell JT, Perazzolli M, Eldredge

G, Gatto P, Oyzerski R, Moretto M, Gutin N, Stefanini M, Chen Y, Segala C, Davenport C, Demattè L, Mraz A, Battilana J, Stormo K, Costa F, Tao Q, Si-Ammour A, Harkins T, Lackey A, Perbost C, Taillon B, Stella A, Solovyev V, Fawcett JA, Sterck L, Vandepoele K, Grando SM, Toppo S, Moser C, Lanchbury J, Bogden R, Skolnick M, Sgaramella V, Bhatnagar SK, Fontana P, Gutin A, Van de Peer Y, Salamini F, Viola R. 2007. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS One **2**:e1326.

Verdier J, Zhao J, Torres-Jerez I, Ge S, Liu C, He X, Mysore KS, Dixon RA, Udvardi MK. 2012. MtPAR MYB transcription factor acts as an on switch for proanthocyanidin biosynthesis in *Medicago truncatula*. Proceedings of the National Academy of Sciences **109**: 1766–1771.

Verweij W, Spelt C, Di Sansebastiano G-P, Vermeer J, Reale L, Ferranti F, Koes R, Quattrocchio F. 2008. An H^+ P-ATPase on the tonoplast determines vacuolar pH and flower colour. Nature Cell Biology **10**: 1456-1462.

Verrier PJ, Bird D, Burla B, Dassa E, Forestier C, Geisler M, Klein M, Kolukisaoglu U, Lee Y, Martinoia E, Murphy A, Rea PA, Samuels L, Schulz B, Spalding EJ, Yazaki K, Theodoulou FL. 2008. Plant ABC proteins--a unified nomenclature and updated inventory. Trends Plant Science **13**: 151-9.

Verriès C, Guiraud JL, Souquet JM, Vialet S, Terrier N, Olle D. 2008. Validation of an extraction method on whole pericarp of grape berry (*Vitis vinifera* L. cv. Shiraz) to study biochemical and molecular aspects of flavan-3-ol synthesis during berry development. Journal of Agricultural and Food Chemistry **56**: 5896-5904.

Vidal S, Cartalade D, Souquet J, Fulcrand H, Cheynier V. 2002. Changes in proanthocyanidin chain-length inwine-like model solutions. Journal of Agricultural Food Chemistry **50**: 2261–2266**.**

Villegas RJA, Kojima M. 1986. Purification and characterization of hydroxycinnamoyl D-glucose quinate hydroxycinnamoyltransferase in the root of sweet potato, Ipomoea batatas Lam. Journal of Biological Chemistry **261**: 8729–8733.

Vitrac X, Bornet A, Vanderlinde R, Valls J, Richard T, Delaunay JC, Mérillon JM, Teissédre PL. 2005. Determination of stilbenes (delta-viniferin, trans-astringin, trans-piceid, cis- and trans-resveratrol, epsilon-viniferin) in Brazilian wines. Journal of Agricultural and Food Chemistry **53**: 5664-9.

Wada L, Ou B. 2002. Antioxidant activity and phenolic content of Oregon caneberries. Journal of Agricultural and Food Chemistry **50**: 3495-3500.

Wagner GJ, Hrazdina G. 1984. Endoplasmic reticulum as a site of phenylpropanoid and flavonoid metabolism in hippeastrum. Plant Physiology **74**: 901–906.

Wajant H, Mundry KW, Pfizenmaier K. 1994. Molecular cloning of hydroxynitrile lyase from Sorghum bicolor (L.). Homologies to serine carboxypeptidases. Plant Molecular Biology **26**:735-46.

Walker AR, Lee E, Bogs J, McDavid DA, Thomas MR, Robinson SP. 2007. White grapes arose through the mutation of two similar and adjacent regulatory genes. Plant Journal **49**: 772-85.

Wang LJ, Huang WD, Liu YP, Zhan JC. 2005. Changes in salicylic and abscisic acid contents during heat treatment and their effect on thermotolerance of grape plants. Russian Journal of Plant Physiology **52**: 516–520.

Wang H, Wang W, Li H, Zhang P, Zhan J, Huang W. 2010. Gene transcript accumulation, tissue and subcellular localization of anthocyanidin synthase (ANS) in developing grape berries. Plant Science **179**: 103–113.

Waterhouse AL, Lamuela-Raventos RM. 1994. The occurrence of piceid, a stilbene glucoside, in grape berries. Phytochemistry **37**: 571-573.

Wei YL, Li JN, Lu J, Tang ZL, Pu DC, Chai YR. 2007. Molecular cloning of *Brassica napus* TRANSPARENT TESTA 2 gene family encoding potential MYB regulatory proteins of proanthocyanidin biosynthesis. Molecular Biology Reports **34**: 105–120.

Weier D, Mittasch J, Strack D, Milkowski C. 2008. The genes BnSCT1 and BnSCT2 from Brassica napus encoding the final enzyme of sinapine biosynthesis-molecular characterization and suppression. Planta **227**: 375–385.

Wen PF, Chen JY, Kong WF, Pan QH, Wan SB, Huang WD. 2005. Salicylic acid induced the expression of phenylalanine ammonia-lyase gene in grape berry. Plant Science **169**: 928–934.

Werner I, Bacher A, Eisenreich W. 1997. Retrobiosynthetic NMR studies with 13C-labeled glucose. Formation of gallic acid in plants and fungi. Journal of Biological Chemistry **272**:25474-82.

Werner RA, Rossmann A, Schwarz C, Bacher A, Schmidt HL, Eisenreich W. 2004. Biosynthesis of gallic acid in *Rhus typhina*: discrimination between alternative pathways from natural oxygen isotope abundance. Phytochemistry **65**: 2809-13.

Wheeler S, Loveys B, Ford C, Davies C. 2009. The relationship between the expression of abscisic acid biosynthesis genes, accumulation of abscisic acid and the promotion of *Vitis vinifera* L. berry ripening by abscisic acid. Australian Journal of Grape Wine Research **15**: 195–204.

Widhalm JR, Gutensohn M, Yoo H, Adebesin F, Qian Y, Guo L, Jaini R, Lynch JH, McCoy RM, Shreve JT, Thimmapuram J, Rhodes D, Morgan JA, Dudareva N. 2015. Identification of a plastidial phenylalanine exporter that influences flux distribution through the phenylalanine biosynthetic network. Nature Communications **6**:8142.

Widhalm JR, Dudareva N. 2015. A familiar ring to it: biosynthesis of plant benzoic acids. Molecular Plant **8**: 83-97.

Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM. 2009. Expansion and diversification of the Populus R2R3-MYB family of transcription factors. Plant Physiology **149**: 981-93.

Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiology **126**: 485-493.

Winkel BS. 2004. Metabolic channeling in plants. Annual Review of Plant Biology **55**: 85-107.

Wu BH, Cao YG, Guan L, Xin HP, Li JH, Li SH. 2014. Genome-wide transcriptional profiles of the berry skin of two red grape cultivars (*Vitis vinifera*) in which anthocyanin synthesis is sunlight-dependent or -independent. PLoS One **9**: e105959.

Wyrepkowski CC, Costa DL, Sinhorin AP, Vilegas W, De Grandis RA, Resende FA, Varanda EA, dos Santos LC. 2014. Characterization and quantification of the compounds of the ethanolic extract from *Caesalpinia ferrea* stem bark and evaluation of their mutagenic activity. Molecules **19**:16039–57.

Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA. 2003. Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. Science **299**: 396–99.

Xie DY, Sharma SB, Dixon RA. 2004. Anthocyanidin reductases from *Medicago truncatula* and *Arabidopsis thaliana*. Archive of Biochemistry and Biophysics **422**: 91–102.

Xu W, Grain D, Le Gourrierec J, Harscoët E, Berger A, Jauvion V, Scagnelli A, Berger N, Bidzinski P, Kelemen Z, Salsac F, Baudry A, Routaboul JM, Lepiniec L, Dubos C. 2013. Regulation of flavonoid biosynthesis involves an unexpected complex transcriptional regulation of TT8 expression, in Arabidopsis. New Phytologist **198**: 59-70.

Xu W, Grain D, Bobet S, Le Gourrierec J, Thévenin J, Kelemen Z, Lepiniec L, Dubos C. 2014. Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR complexes and their targets in Arabidopsis seed. New Phytologist **202**: 132–44.

Yamane T, Jeong ST, Goto-Yamamoto N, Koshita Y, Kobayashi S. 2006. Effects of temperature on anthocyanin biosynthesis in grape berry skins. American Journal of Enology and Viticulture **57**: 54–59.

Yang Q, Reinhard K, Schiltz E, Matern U. 1997. Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA:anthranilate N-hydroxycinnamoyl/benzoyltransferase from elicited cell cultures of carnation, Dianthus caryophyllus L. Plant Molecular Biology **35**:777-789.

Yang H, Ye X, Liu D, Chen J, Zhang J, Shen Y, Yu D. 2011. Characterization of unusual proanthocyanidins in leaves of bayberry (*Myrica rubra* Sieb. et Zucc.). Journal of Agricultural and Food Chemistry **59**: 1622–9.

Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q, Yunping S, Li Z, Xiaohui D, Jingchu L, Xing-Wang D, Zhangliang C, Hongya G, Li-Jia Q. 2006. The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Molecular Biology **60**: 107–24.

Yilmaz Y, Toledo RT. 2004. Major flavonoids in grape seeds and skins: antioxidant capacity of catechin, epicatechin, and gallic acid. Journal of Agricultural and Food Chemistry **52**: 255–60.

Yonekura-Sakakibara K, Fukushima A, Nakabayashi R, Hanada K, Matsuda F, Sugawara S, Inoue E, Kuromiri T, Ito T, Shinozaki K, Wangwattana B, Yamazaki M, Saito K. 2012. Two glycosyltransferases involved in anthocyanin modification delineated by transcriptome independent component analysis in *Arabidopsis thaliana*. Plant Journal **69**: 154–167.

Yoo H, Widhalm JR, Qian Y, Maeda H, Cooper BR, Jannasch AS, Gonda I, Lewinsohn E, Rhodes D, Dudareva N. 2013. An alternative pathway contributes to phenylalanine biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase. Nature Communications **4**: 2833.

Yoshida K, Iwasaka R, Kaneko T, Sato S, Tabata S, Sakuta M. 2008. Functional differentiation of *Lotus japonicus* TT2s, R2R3-MYB transcription factors comprising a multigene family. Plant Cell Physiology **49**: 157–169.

Yoshida K, Oyama KI, Kondo T. 2012. Chemistry of flavonoids in color development, in: Cheynier V, Sarni-Manchado P, Quideau S (Eds.), Recent Advances in Polyphenols Research, vol. 3, Wiley-Blackwell Publishing, Oxford, UK, 2012, pp. 99e129.

Zarena AS, Sankar KU. 2012. Phenolic acids, flavonoid profile and antioxidant activity in mangosteen (*Garcinia mangostana* L.) pericarp. Journal of Food Biochemistry **36**: 627–633.

Zhang X, Luo GG, Wang RH, Wang J, Himelrick DG. 2003. Growth and developmental responses of seeded and seedless grape berries to shoot girdling. Journal of the American Society for Horticultural Science **128**: 316–323.

Zhang XY, Wang XL, Wang XF, Xia GH, Pan QH, Fan RC, Wu FQ, Yu XC, Zhang DP. 2006. A shift of phloem unloading from symplasmic to apoplasmic pathway is involved in developmental onset of ripening in grape berry. Plant Physiology **142**: 220-232.

Zhao J, Dixon RA. 2009. MATE transporters facilitate vacuolar uptake of epicatechin 3'-*O*-glucoside for proanthocyanidin biosynthesis in *Medicago truncatula* and Arabidopsis. Plant Cell **21**: 2323-2340.

Zhao J, Dixon RA. 2010. The "ins" and "outs of flavonoid transport. Trends Plant Science **15**: 72–80.

Zhao J, Pang Y, Dixon RA. 2010. The mysteries of proanthocyanidin transport and polymerization. Plant Physiology **153**: 437-43.

Zhao J. 2015. Flavonoid transport mechanisms: how to go, and with whom. Trends in Plant Science **20**:576-85.

Zhou A, Li J. 2005. Arabidopsis BRS1 is a secreted and active serine carboxypeptidase. Journal of Biological Chemistry **280**: 35554–61.

Zifkin M, Jin A, Ozga JA, Zaharia LI, Schernthaner JP, Gesell A, Abrams SR, Kennedy JA, Constabel CP. 2012. Gene expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism and activation of abscisic acid metabolism. Plant Physiology **158**: 200–224.

Ziliotto F, Corso M, Rizzini F, Rasori A, Botton A, Bonghi C. 2012. Grape berry ripening delay induced by a pre-véraison NAA treatment is paralleled by a shift in the expression pattern of auxin-and ethylene-related genes. BMC Plant Biology **12**: 185.

ANNEXES

Culture media

Microorganism culture

LB

Tryptone (Bactotryptone)	10 $g.L^{-1}$
Yeast extract (YE)	5 $g.L^{-1}$
NaCl	10 $g.L^{-1}$

Luria Bertani (LB) was used for the growth of *E. coli* and *A.tumefaciens*strainsin liquid or on agar medium (agar 15$g.L^{-1}$). It is sterilized by autoclaving (20 min, 120°C) then made selective by the addition of suitable antibiotics (ampicillin 100 $\mu g.mL^{-1}$, kanamycin 50 $\mu g.mL^{-1}$and spectinomycin75 $\mu g.mL^{-1}$).

YTA2X

Tryptone (Bactotryptone)	10 $g.L^{-1}$
Yeast extract (YE)	10 $g.L^{-1}$
NaCl	5 $g.L^{-1}$

Autoclaved (20 min, 120°C).

MGL/B

Mannitol ($C_6H_{14}O_6$)	5 $g.L^{-1}$
L-glutamate ($C_5H_8NO_4Na$)	1 $g.L^{-1}$
Potassium phosphate (KH_2PO_4)	0.15 $g.L^{-1}$
Magnesium sulphate ($MgSO_4$, $7H_2O$)	0.1 $g.L^{-1}$
Iron EDTA ($FeSO_4$, $7H_2O$)	2.5 $mL.L^{-1}$
Tryptone (Bactotryptone)	5 $g.L^{-1}$
Yeast extract	2.5 $g.L^{-1}$
NaCl	5 $g.L^{-1}$
Biotin 100X	5 $mL.L^{-1}$

MGL/B medium was used for the growth of *Agrobacterium rhizogenes* strains in liquid medium or on agar medium (Micro Agar 10 $g.L^{-1}$). pH was adjusted to 7 with a NaOH solution. MGL/B medium was sterilized by autoclaving (20 min, 120°C) and then made selective by the addition of suitable antibiotics.

Plant material culture

MS/2

Macro MS (10X)	50 $mL.L^{-1}$
Micro MS (100X)	10 $mL.L^{-1}$
Iron (EDTA) (200X)	5 $mL.L^{-1}$
Vitamin MG	1 $mL.L^{-1}$
Sucrose	20 $g.L^{-1}$

The medium MS/2 was used for growing grapevine vitroplants on agar medium (Micro Agar 8 $g.L^{-1}$). The pH was adjusted to 6.5 with a KOH solution and then sterilized by autoclaving (20 min, 120°C).

Macro MS (10X)

KNO_3	19 $g.L^{-1}$
NH_4NO_3	16.5 $g.L^{-1}$
$CaCl_2$, $2H_2O$	4.4 $g.L^{-1}$
$MgSO_4$, $7H_2O$	3.7 $g.L^{-1}$
KH_2PO_4	1.7 $g.L^{-1}$

Not autoclaved and stored at 4°C.

MicroMS(100X)

$MnSO_4$, H_2O	2.23 $g.L^{-1}$
H_3BO_3	0.12 $g.L^{-1}$
$ZnSO_4$, $7H_2O$	0.36 $g.L^{-1}$
Na_2MoO_4, $2H_2O$	25 $mg.L^{-1}$
$CuSO_4$, $5H_2O$	2.5 $mg.L^{-1}$
Ki	83 $mg.L^{-1}$
$CoCl_2$, $6H_2O$	25 $mg.L^{-1}$

Not autoclaved

Iron (EDTA. 200X)

$FeSO_4$, $7H_2O$ 99%	4.552 $g.L^{-1}$
Na (EDTA)	7.45 $g.L^{-1}$

Sterile filtered and stored in the dark at 4°C.

Vitamin MG

Myo inositol ($C_6H_{12}O_6$)	50 $g.L^{-1}$
Nicotinic acid ($C_6H_5NO_2$)	1 $g.L^{-1}$
PyrodoxineHCl	1 $g.L^{-1}$
Thiamine HCl	1 $g.L^{-1}$
Calcium pantothenate	1 $g.L^{-1}$

Prepared under hood.

LGO

Macro LGO	100 $mL.L^{-1}$
Micro MS	10 $mL.L^{-1}$
Iron (EDTA)	10 $mL.L^{-1}$
ReinforcedVitamin	2 $mL.L^{-1}$
Casein	0.25 $g.L^{-1}$
Sucrose	25 $g.L^{-1}$

LGO medium was used for the growth of hairy roots on agar medium (Phytagel 5 $g.L^{-1}$). The pH was adjusted to 6 with a KOH solution and then sterilized by autoclaving (20 min. 120°C).To avoidcontamination of the culture medium, cefotaxime (200 $\mu g.mL^{-1}$) and augmentin (200 $\mu g.mL^{-1}$) were added to the medium.

Macro LGO

KNO_3	15 $g.L^{-1}$
NH_4NO_3	1.5 $g.L^{-1}$
$CaCl_2$, $2H_2O$	1.5 $g.L^{-1}$
$MgSO_4$, $7H_2O$	2.5 $g.L^{-1}$
NaH_2PO_4, H_2O	2.5 $g.L^{-1}$

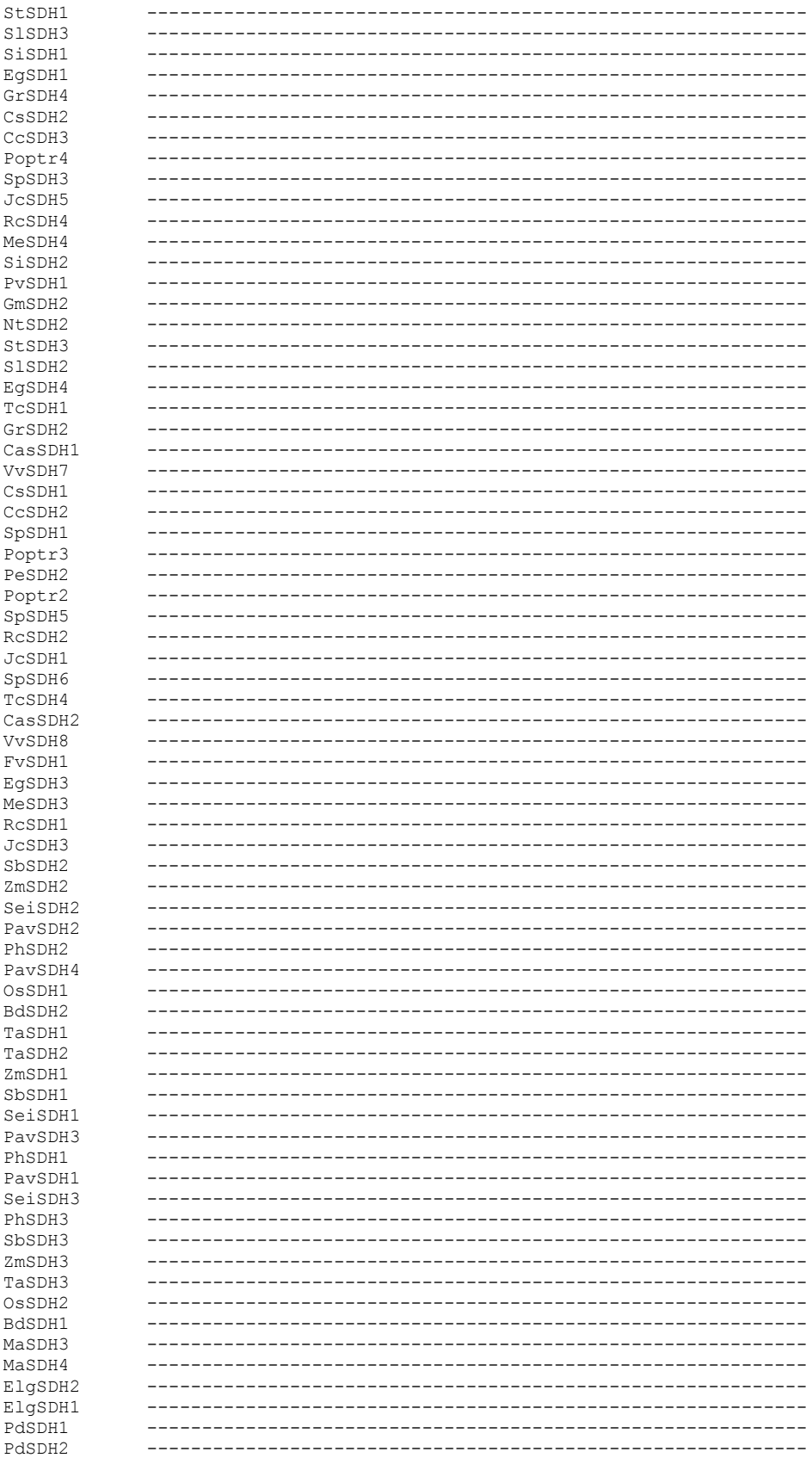
Not autoclaved and stored at 4 ° C.

Reinforced vitamins

Myo inositol ($C_6H_{12}O_6$)	10 $mg.L^{-1}$
Nicotinic acid ($C_6H_5NO_2$)	10 $mg.L^{-1}$
Thiamine HCl	10 $mg.L^{-1}$
Pyridoxine HCl	1 $mg.L^{-1}$
Calcium pantothenate	1 $mg.L^{-1}$
Biotin 100X	0.5 $mL.L^{-1}$

Prepared under hood.

Supplemental Figure 1.



```
MaSDH1      ------------------------------------------------------------
MaSDH2      ------------------------------------------------------------
PvSDH2      ------------------------------------------------------------
GmSDH1      ------------------------------------------------------------
FvSDH2      ------------------------------------------------------------
PpSDH1      ------------------------------------------------------------
PmSDH3      ------------------------------------------------------------
MdSDH1      ------------------------------------------------------------
PbSDH1      ------------------------------------------------------------
MdSDH4      ------------------------------------------------------------
PbSDH4      ------------------------------------------------------------
TcSDH3      ------------------------------------------------------------
GrSDH3      ------------------------------------------------------------
CasSDH3     ------------------------------------------------------------
DkSDH       ------------------------------------------------------------
EgSDH2      ------------------------------------------------------------
VvSDH9      ------------------------------------------------------------
SpSDH2      ------------------------------------------------------------
Poptr5      ------------------------------------------------------------
PeSDH3      ------------------------------------------------------------
MeSDH1      ------------------------------------------------------------
JcSDH4      ------------------------------------------------------------
CpSDH       ------------------------------------------------------------
AtSDH       ------------------------------------------------------------
AlSDH       ------------------------------------------------------------
PpSDH2      ------------------------------------------------------------
PmSDH2      ------------------------------------------------------------
PmSDH1      ------------------------------------------------------------
MdSDH2      ------------------------------------------------------------
PbSDH2      ------------------------------------------------------------
MdSDH5      MLXFPGLAPGFRLHQWMLWXFAVVWPRYSTVXTAARGHIVSSILGPYPVHSPVATSXXFL
FvSDH3      ------------------------------------------------------------
FsSDH       ------------------------------------------------------------
JrSDH       ------------------------------------------------------------
FvSDH4      ------------------------------------------------------------
PpSDH3      ------------------------------------------------------------
MdSDH3      ------------------------------------------------------------
PbSDH3      ------------------------------------------------------------
NnSDH1      ------------------------------------------------------------
NnSDH2      ------------------------------------------------------------
SiSDH3      ------------------------------------------------------------
NtSDH1      ------------------------------------------------------------
StSDH2      ------------------------------------------------------------
SlSDH1      ------------------------------------------------------------
MtSDH       ------------------------------------------------------------
PvSDH3      ------------------------------------------------------------
GmSDH3      ------------------------------------------------------------
LuSDH1      ------------------------------------------------------------
LuSDH2      ------------------------------------------------------------
TcSDH2      ------------------------------------------------------------
GrSDH1      ------------------------------------------------------------
SpSDH4      ------------------------------------------------------------
Poptr1      ------------------------------------------------------------
PeSDH1      ------------------------------------------------------------
MeSDH2      ------------------------------------------------------------
RcSDH3      ------------------------------------------------------------
JcSDH2      ------------------------------------------------------------
VvSDH5      ------------------------------------------------------------
EgSDH5      ------------------------------------------------------------
CsSDH3      ------------------------------------------------------------
CcSDH1      ------------------------------------------------------------
```

```
StSDH1      ------------------------------------------------------------
SlSDH3      ----------------------------------------------MF----PPLY--
SiSDH1      ------------------------------------------------------------
EgSDH1      --------------------MA--------KLHATPQAS-RAGE-SKRE----YPIH--
GrSDH4      ------------------------------------------------------------
CsSDH2      ------------------------------------------------------------
CcSDH3      ------------------------------------------------------------
Poptr4      ------------------------------------------------------------
SpSDH3      ------------------------------------------------------------
JcSDH5      ------------------------------------------------------------
RcSDH4      ------------------------------------------------------------
MeSDH4      ------------------------------------------------------------
SiSDH2      ------------------------------------------------------------
PvSDH1      ------------------------------------------------------------
GmSDH2      ------------------------------------------------------------
NtSDH2      ------------------------------------------------------------
StSDH3      ------------------------------------------------------------
SlSDH2      ------------------------------------------------------------
EgSDH4      ------------------------------------------------------------
TcSDH1      ------------------------------------------------------------
GrSDH2      ------------------------------------------------------------
CasSDH1     ------------------------------------------------------------
VvSDH7      ------------------------------------------------------------
CsSDH1      ------------------------------------------------------------
CcSDH2      ------------------------------------------------------------
SpSDH1      ------------------------------------------------------------
Poptr3      ------------------------------------------------------------
PeSDH2      ------------------------------------------------------------
Poptr2      ------------------------------------------------------------
SpSDH5      ------------------------------------------------------------
RcSDH2      ------------------------------------------------------------
JcSDH1      ------------------------------------------------------------
SpSDH6      ------------------------------------------------------------
TcSDH4      ------------------------------------------------------------
CasSDH2     ------------------------------------------------------------
VvSDH8      ------------------------------------------------------------
FvSDH1      ------------------------------------------------------------
EgSDH3      ------------------------------------------------------------
MeSDH3      -------------------------MLNTFKLA---TIKSLT------------TLLRT
RcSDH1      ------------------------------------------------------------
JcSDH3      ------------------------------------------------------------
SbSDH2      ------------------------------------------------------------
ZmSDH2      ------------------------------------------------------------
SeiSDH2     ------------------------------------------------------------
PavSDH2     ------------------------------------------------------------
PhSDH2      --------------------MQGLC-------LLGPERGPSATPAHLHSPPAAPQVHHQ
PavSDH4     ------------------------------------------------------------
OsSDH1      ------------------------------------------------------------
BdSDH2      ------------------------------------------------------------
TaSDH1      ------------------------------------------------------------
TaSDH2      ------------------------------------------------------------
ZmSDH1      ------------------------------------------------------------
SbSDH1      ------------------------------------------------------------
SeiSDH1     ------------------------------------------------------------
PavSDH3     ------------------------------------------------------------
PhSDH1      ------------------------------------------------------------
PavSDH1     ------------------------------------------------------------
SeiSDH3     ------------------------------------------------------------
PhSDH3      ------------------------------------------------------------
SbSDH3      ------------------------------------------------------------
ZmSDH3      ------------------------------------------------------------
TaSDH3      ------------------------------------------------------------
OsSDH2      ------------------------------------------------------------
BdSDH1      ------------------------------------------------------------
MaSDH3      ------------------------------------------------------------
MaSDH4      ----------------------------------------------------------MNHQ
ElgSDH2     ------------------------------------------------------------
ElgSDH1     ------------------------------------------------------------
PdSDH1      ------------------------------------------------------------
PdSDH2      ------------------------------------------------------------
MaSDH1      ------------------------------------------------------------
MaSDH2      ------------------------------------------------------------
PvSDH2      ------------------------------------------------------------
GmSDH1      ------------------------------------------------------------
```

```
FvSDH2      ------------------------------------------------------------
PpSDH1      ------------------------------------------------------------
PmSDH3      ------------------------------------------------------------
MdSDH1      ------------------------------------------------------------
PbSDH1      ------------------------------------------------------------
MdSDH4      ------------------------------------------------------------
PbSDH4      ------------------------------------------------------------
TcSDH3      ------------------------------------------------------------
GrSDH3      ------------------------------------------------------------
CasSDH3     ------------------------------------------------------------
DkSDH       ------------------------------------------------------------
EgSDH2      ------------------------------------------------------------
VvSDH9      ------------------------------------------------------------
SpSDH2      ------------------------------------------------------------
Poptr5      ------------------------------------------------------------
PeSDH3      ------------------------------------------------------------
MeSDH1      ------------------------------------------------------------
JcSDH4      ------------------------------------------------------------
CpSDH       ------------------------------------------------------------
AtSDH       ---------------MAASSTNA--RLTNPPRLLSKPRLSPTSVA-NLRFPAADFSTRFF
AlSDH       ---------------MAASSTNA--RLTNPPHVLSKPRLSPTSVV-NLRFPAADFPIRLS
PpSDH2      ------------------------------------------------------------
PmSDH2      ------------------------------------------------------------
PmSDH1      ------------------------------------------------------------
MdSDH2      ------------------------------------------------------------
PbSDH2      ------------------------------------------------------------
MdSDH5      GHXXRSSLPISMVSTLASVEGXATLVVLNSHRIXKTPT-KPLCGA-THRTPKA-ASLRRP
FvSDH3      ------------------------------------------------------------
FsSDH       ------------------------------------------------------------
JrSDH       ------------------------------------------------------------
FvSDH4      ------------------------------------------------------------
PpSDH3      ------------------------------------------------------------
MdSDH3      ----MAKIPTKL-------------------IPLLTPT-TPLRAA-SHQPQPQHQKQQRL
PbSDH3      ------------------------------------------------------------
NnSDH1      ------------------------------------------------------------
NnSDH2      ------------------------------------------------------------
SiSDH3      ------------------------------------------------------------
NtSDH1      ------------------------------------------------------------
StSDH2      ------------------------------------------------------------
SlSDH1      ------------------------------------------------------------
MtSDH       ------------------------------------------------------------
PvSDH3      ------------------------------------------------------------
GmSDH3      ------------------------------------------------------------
LuSDH1      ------------------------------------------------------------
LuSDH2      ------------------------------------------------------------
TcSDH2      ------------------------------------------------------------
GrSDH1      ------------------------------------------------------------
SpSDH4      ------------------------------------------------------------
Poptr1      ------------------------------------------------------------
PeSDH1      ------------------------------------------------------------
MeSDH2      ------------------------------------------------------------
RcSDH3      ------------------------------------------------------------
JcSDH2      ------------------------------------------------------------
VvSDH5      ------------------------------------------------------------
EgSDH5      ------------------------------------------------------------
CsSDH3      ------------------------------------------------------------
CcSDH1      ------------------------------------------------------------
```

```
StSDH1      ------------------------------------------------------MGLKHDL
SlSDH3      -----KRERWS-------------------------------VKEERDIIMGLKNDL
SiSDH1      ------------------------------------------------------------
EgSDH1      -----RERSWTVSPPDHWRKQVAEGV--TAQTLHSSSRIFFLHCPSSRRERPREMDRNGV
GrSDH4      ------------------------------------------------------MDVKNDV
CsSDH2      -----------------------------------------------------MEVAAKNSL
CcSDH3      ------------------------------------------------------------
Poptr4      ------------------------------------------------------MAFKNNL
SpSDH3      ------------------------------------------------------MAFKNSI
JcSDH5      ------------------------------------------------------MALQNDV
RcSDH4      ------------------------------------------------------MAFQNNL
MeSDH4      ------------------------------------------------------------
SiSDH2      ---------------------------------------------MGRAGF-SKNSR
PvSDH1      ------------------------------------------------------------
GmSDH2      ---------------------------------------------MGSNQV-WKHSV
NtSDH2      ---------------------------------------------MGSVGL-LKNSA
StSDH3      ---------------------------------------------MGSVGL-LKNSA
SlSDH2      ---------------------------------------------MGSVGL-LKNSA
EgSDH4      ---------------------------------------------MGSLSLSSVGLT
TcSDH1      ---------------------------------------------MGGTGM-LNNST
GrSDH2      ---------------------------------------------MGSTGI-LKNST
CasSDH1     ---------------------------------------------MGSVGV-LTNST
VvSDH7      ---------------------------------------------MDDVGV-LKKET
CsSDH1      ---------------------------------------------MGVVNI-TKNTT
CcSDH2      ---------------------------------------------MGVVNI-TKNTT
SpSDH1      ---------------------------------------------MGSTAGA-FKNSA
Poptr3      ---------------------------------------------MGSVGV-LTNSV
PeSDH2      ---------------------------------------------MGSVGV-LTNSV
Poptr2      ---------------------------------------------MGRAGI-LANST
SpSDH5      ---------------------------------------------MERARI-LTNST
RcSDH2      ---------------------------------------------MGSVGE-LTSSV
JcSDH1      ---------------------------------------------MGHVGL-LNSSA
SpSDH6      ----------------------------------MESLQFM-TSGSLQTSTNEVRGSPT
TcSDH4      ------------------------------------------------MSHSRVQSRPT
CasSDH2     --------------------------------MASGSFSFA--TSDVQTSSSGVRSGPT
VvSDH8      ----------------------------------MGSLPFT--VSDLQTSVSGVRSNPT
FvSDH1      ----------------------------------MGSLPFT--TSDLHTSTGGFLSSPT
EgSDH3      ----------------------------------MGSVPFT--TSDLQTSTSGFRSSPT
MeSDH3      VTRLPGVLRPAPPP----------------PSSSAMGSLPFT--MSDLQLSSKEVQGSPT
RcSDH1      ---------------------------------------------MSDLQLSNSSVQSSPT
JcSDH3      ----------------------------------MGSLRFT--MSDLQLSNREVQSSPT
SbSDH2      --------------------------------MATAMS--V-SAVT-PA-AAAVARPRT
ZmSDH2      --------------------------------MATAMS--V-SAVS-PP-AAAVTRPRT
SeiSDH2     --------------------------------MATAMS--V-SAVS-PA-AAAVARPRT
PavSDH2     ------------------------------------------------------------
PhSDH2      PRVQCRRQLNPLLP-LPLPSPTT--------A-MATAMS--V-SAVS-PA-ATAVARPRT
PavSDH4     --------------------------------MATAMS--V-SSAS-PAAAAAVARPRT
OsSDH1      ------------------------------------MS--A-SAMVPPSAAAAVARART
BdSDH2      ------------------------------------MS--V-SAVVPPA--AAVARART
TaSDH1      ------------------------------------------------------------
TaSDH2      ------------------------------------------------------------
ZmSDH1      -----------------MRVPPG--------S-HA-----------YTSAAELNAPLRGK
SbSDH1      --------------------------------MA-----------TAATSTP-PAWQRT
SeiSDH1     ---------------------------------------------MERSARRGT
PavSDH3     --------------------------------MA-----------AAAASAPAAPRGTT
PhSDH1      --------------------------------MA-----------TAAASAPAAPLG-T
PavSDH1     --------------------------------MA-----------AAAASAPAAPRG-T
SeiSDH3     ------------------------------------------------------MT
PhSDH3      ------------------------------------------------------MT
SbSDH3      ------------------------------------------------------MT
ZmSDH3      ------------------------------------------------------MT
TaSDH3      ------------------------------------------------------------
OsSDH2      ------------------------------------------------------MT
BdSDH1      ------------------------------------------------------MT
MaSDH3      ------------------------------------------------------------
MaSDH4      IRYAVATLTWILIS-SSLYKL----------H-RRRAFS--L-----RLRYQIQRGRGMT
ElgSDH2     ---------------------------------------------MAMEGST
ElgSDH1     ---------------------------------------------MAMEGST
PdSDH1      ------------------------------------------------------------
PdSDH2      ---------------------------------------------MAMEGST
MaSDH1      ------------------------------------------------------MT
MaSDH2      ------------------------------------------------------MT
PvSDH2      --------------------------------MTTTTS--A-PFQWD-------RRNST
GmSDH1      --------------------------------MSISS--A-PFQWEAGIGSGVRRNST
FvSDH2      --------------------------------MTLSSI--P-LVASDLQISYGTGRNST
```

```
PpSDH1      ------------------------------------------------------------
PmSDH3      ----------------------------------------M-SAVSDLHTTNGTRRNST
MdSDH1      --------------------------------MTRSTL--P-LAASDLHIMDGTRRNST
PbSDH1      --------------------------------MTRGTL--P-LAASDLHIMDGTRRNST
MdSDH4      ------------------------------------------------------------
PbSDH4      --------------------------------MTLDTL--P-LAASDLHIMDGSRRNST
TcSDH3      --------------------------------MAFESS--T-LPITDFQIGDGAQRNST
GrSDH3      --------------------------------MALESF--Q-LPITDFHIGNETRRNSV
CasSDH3     ------------------------------------------------------------
DkSDH       ------------------------------------------------------------
EgSDH2      --------------------------------MTLSSI--P-LTAADLQIPAGGRRNST
VvSDH9      --------------------------------MTLSSV--P-LATSDIQIPEGARRNST
SpSDH2      --------------------------------MTLGRV--S-LTSMDLQSIDEMRRNST
Poptr5      ------------------------------------------------MDLQSADGERRNST
PeSDH3      ------------------------------------------------MDLQSADGVRRNST
MeSDH1      --------------------------------MTLGSV--P-LTTTDIQIVDGAQRNST
JcSDH4      --------------------------------MTPGSV--P-LTTLDLRIEEGARRNST
CpSDH       ------------------------------------------------MEGGDISRNAT
AtSDH       ADSSSPRLRSVPFP-VVFSDQRRRRSMEPSNVYVA---S--N-STEMEIGSHD-IVKNPS
AlSDH       AGSSSPQLRSVPFP-VVFSDRSRRRTMEHSNVYVA---S--N-STDMEIGSQD-IVKNPS
PpSDH2      ------------------------------------------------------------
PmSDH2      --------------------------------MLS---S--C-SDALKVGGEGMMRRNST
PmSDH1      --------------------------------MLS---S--C-SDALKVGGEGMMRRNST
MdSDH2      MDLTDSLLFWLNF-------------DQLDSSSSA---A--G-TDGLEVKGNGMRARNST
PbSDH2      --------MQSHH-------------ITLDSSSSA---A--G-SDGLEVKCNGMRARNST
MdSDH5      S--------NSPM-------------DSPQNV-LL--------ASEGRVGVED-VRRSST
FvSDH3      ------------------------------------------------------MMGSST
FsSDH       ------------M-------------DSTNVL-LA---S--S-SASLQIGGGE-IRKNPT
JrSDH       ------------------------------------------------------------T
FvSDH4      ------------------------------MDS-PT---L--T-VASAQVGGGG-MRKSST
PpSDH3      ------------M-------------DSTQNI-LL---A--S-SGGLQVGSKG-VRRSST
MdSDH3      SGHXPQTLRWILP-------------KTFWKI-YK---L--A-SDGLQVGVDG-VRRSST
PbSDH3      ----------------------------MDSPQN-VL---L--A-SDGLQVGVDG-VRRSST
NnSDH1      ------------------------------------------------------------
NnSDH2      ----------------------------------------------------MAGNST
SiSDH3      ------------------------------------------------MGSGVEMRKNQT
NtSDH1      ----------------------------------------MELV--V-DSGVRKMEGEAMTRNET
StSDH2      ----------------------------------------------MEGEAMRKNQT
SlSDH1      ----------------------------------MELV--V-DSGVKKMEGEAMRKNQT
MtSDH       ------------------------------------------------------MNAM
PvSDH3      --------------------------MDSPNV---------------LFTAPSAGT-KNAT
GmSDH3      --------------------------MDSPNA---------------LSTAPAAGSRKNAT
LuSDH1      --------------------------MDSGNVMLTSSSSA--M------GGGTGLPLNNPT
LuSDH2      --------------------------MDSGNVMLTSSSSA--M------GGGTGLPLNNPT
TcSDH2      --------------------------MDTLNPLLVSSSAG--L-NM---ESRGG-VWKNPT
GrSDH1      --------------------------MDTVNPLLVSSASG--L-KM---DVNGGEMRKNPT
SpSDH4      --------------------------MDSASNDLLVSSPS--A-AAVV-GMASGGARKNPT
Poptr1      --------------------------MDSASNVLLASSPS--A-AAAV-GMGSGGVRRNPT
PeSDH1      --------------------------MDSASNVLLASSPS--A-AAVV-GMGSGAVRRNPT
MeSDH2      --------------------------MDSGGSVLLASSST--A-AEQIGGGGVGGMRKNPT
RcSDH3      --------------------------MDSTGSALLASSTS--T-DAEMGSRGVGGLRKNPT
JcSDH2      --------------------------MDSTGNVLLASSST--A-VVQMGAGGVGGMRKNPT
VvSDH5      ----------------------------------------------MESGGMSKNST
EgSDH5      ----------------------------------------------MESGAVRKNST
CsSDH3      --------------------------ME--------SPNL--L-VASGSKLVSGGMRKNPT
CcSDH1      --------------------------ME--------SPNL--L-VASGSKLVSGGMRKNPT
```

```
StSDH1       VVYTRLEC---ESCEEMAASMHKAKEEGADLVELCI------DDFTFSHI--SQLELLL-
SlSDH3       VVYTRLEC---ESCEEMTCCIEKAKEEGADLVELCI------DDFTFSDI--SQLEELL-
SiSDH1       ------------------------------------------------------------
EgSDH1       LVCAPLEC---ETLEGMLSSMDKAKAHGADLVELRV------DAMSFGRV--SEVEELI-
GrSDH4       VVCTPLEF---ESLEEMVASMEKAKAEGADVVELCM------ASMAAPYLSHQVVDQLL-
CsSDH2       LVCTQLEC---ETTEEMQASIEQAKVEGADLVELCI------DSMEFSHI--SEVDKLI-
CcSDH3       ------------------------------------------------------------
Poptr4       LVCTPLEC---ETAGEMLSSMKRAETEGADLTELRL------DSLSFSHN--SEVEKLI-
SpSDH3       SVCTPLEC---ESVGEMLASMKRAEAEGADLAELRL------DSLSFSHD--SEVEKLI-
JcSDH5       VVCTPLEC---ETAREMLSSMEKARAEGADLVELHI------DSMFFSHI--SQAENLF-
RcSDH4       LVCTPLEC---ETAGEMLSCMEKARGEGADLVELRV------DSMSFSHI--SQVEKLI-
MeSDH4       ------------------------------------------------------------
SiSDH2       MICAPLMA---ESVEQMLSDMHQAKMEGADVVEIRL------GCISNFQ-PHRDLKLLL-
PvSDH1       --------------------MKHAKTEGADAVEVWL------DCITNFH-PLLHLKIIL-
GmSDH2       MVCAAITTHQYVSAEQIVNGMHQAKAEGADIVELRL------DCITNFH-SHHDLKIIL-
NtSDH2       MVCAPLMA---PSVEQLVHGMLQAKAQGADLVEIRL------DGINNFQ-PQKDLQVLL-
StSDH3       MVCAPLMS---TSVDQLIDDMVEAKSQGADCVEIRL------HAIHNFL-PHKHLQLLF-
SlSDH2       MVCAPLMA---TSVDQLIDEMVEAKSQGADCVEIRL------DAIHNFQ-PHKHLQLLF-
EgSDH4       MVCAPVMG---ESVDQVVEEMHKAKAQGADVVEVRL------DCIKQFQ-AHQDLEIIL-
TcSDH1       MICAPLMA---QSVEQMVKDMHQAKAEGAQLVEIRL------DYIKNFQ-PHQDIQIIL-
GrSDH2       MICAPLMG---ETVEQMVKEMKQAKAEGSHLLEIRL------DHITGFH-PHHHLPAIL-
CasSDH1      TICAPLMS---QSVEQMVSDMNQAKAQGADVVEIRL------DCIKDFQ-PNRDLQILL-
VvSDH7       MICTPLMG---QSVEQMVRDMHKAKVEGADLVEVRL------DYINNFH-PQQDLEIIL-
CsSDH1       MICAPLMA---QSVEQVLSNMYQAKAEGADVVEIRL------DCINNFQ-PGKDLEIIL-
CcSDH2       MICAPLMA---QSVEQVLSNMYQAKAEGADVVEIRL------DCINNFQ-PGKDLEIIL-
SpSDH1       MVCAPLMA---RSAEQMVIDMHSAKAQGADVVEVRL------DCIHKFQ-PGQDLETII-
Poptr3       MVCAPLMA---QSVEQMVIDMHSAKAQGADVVEVRL------DCISKFQ-PRQDLETII-
PeSDH2       MVCAPLMA---QSVEQMVLDMHSAKAQGADVVEVRL------DCINKFQ-PSQDLETII-
Poptr2       MVCAPLMA---RSVEQMVIDMQSAEAQGADAVEVRL------DYINSFQ-PSQDLETII-
SpSDH5       VVCAPLMA---RSVEQMVIDMQRAKAQGADAVEVRL------DFIDSFQ-PSQDLETII-
RcSDH2       MVCTPLMA---QSVEQMISDMYNAKTQGADVVEVRL------DYIDNFQ-PPQDLQAIL-
JcSDH1       MVCAPLMA---QSVEQMVRDMYSAKSQGADVVEVRL------DYIKNFQ-PGQDLETIL-
SpSDH6       LLCTPLRG---TTVDQMLMQMRKAKNIGADVVEIRL------ARMSD-------------
TcSDH4       LLCTPITG---TTVGQMLTEMRLAKEIGSDVVEIRL------DCLRNF-SPRQDLEILI-
CasSDH2      LLCTPLIG---TTVDQMLTDMRKAKEIGADMVEIRL------DCLREF-NPRPDLQILI-
VvSDH8       LLCTPLMG---TTVEQMLTEMRKAKEIGADIVEIRL------DCLRNF-SPAQDLQILI-
FvSDH1       LLCTPLMG---TTVDQMLIEMHKAKEIGSDVVEIRL------DCLRNF-NPSSDLQILI-
EgSDH3       LLCTPLMG---TTVDQMLIEMRKAKEIGADLVEIRL------DCLRNF-NPHQDLQILI-
MeSDH3       LLCTPLMG---TAVDQMLIEMQRAKDIGADVVEIRL------DCLKNF-SPRQDLETLI-
RcSDH1       LLCTPLMG---TTVDQMLLEMRKATEISADVVEIRL------DCLRNL-NPRQDLEILI-
JcSDH3       LLCTPLMG---TTVDQMLIEMQKAKEIGADVVEIRL------DCLRNF-SPTQDLEILI-
SbSDH2       MVCVPATA---RAPREMAEELAAAAALGADVAELRL------DCLAGF-APRRDLPVILA
ZmSDH2       MVCVPATA---RAPREMAAELAAAAALGADVAELRL------DCLAGF-APRRDLPVILA
SeiSDH2      LVCVPAMA---RAPREMVAELATAAALGADVAELRL------DRLAGF-APRRDLPVILA
PavSDH2      ----------------MAAELAAAAALGADVAELRL------DCLAGF-APRRDLPVILA
PhSDH2       LVCVPATA---RAPREMAAELAAAAALGADVAELRL------DLLAGF-APRRDLPIILA
PavSDH4      LVCVPATA---RAPREMAAELAAAAALGADLAELRL------DRLAGF-APRRDLPVILA
OsSDH1       LLCVPATA---RAPREMAAELAAAAALGADVAELRL------DRLAGF-APRRDLPVLLA
BdSDH2       LLCVPATA---RAPREMAAEVAAAAALGADVAELRL------DRLSGF-APRRDLPALLA
TaSDH1       ------------------------------------------------------------
TaSDH2       ------------------------------------------------------------
ZmSDH1       LICCSQVP---PATAPS----------ANGDCSDFYAGGVAVYSGLLWS-TPCRDLPVLLA
SbSDH1       LVCASLTA---RVPREMAAEVDAAAALGADVAELQV------GCLDGF-EPRRDLPVLLA
SeiSDH1      LVCASLTA---RSPREMAAEVAAAAALGADVAELRV------DCLDGF-QPRMDLPVLLA
PavSDH3      LVCASLTA---RSPQEMAAEVAAAAALGADVVELQV------GCLEGF-QPRRDLPVLLA
PhSDH1       LVCASLTA---RSPQEMAAEVAATAALGADVAELQV------GCLDGF-QPRRDLPVLLA
PavSDH1      LVCASLTA---RSPQEVAAEGAAAAALGADVVELLV------GCLEGF-QPRRDLPVLLA
SeiSDH3      LICVPLVA---RTVEEMVADAVAAAAAGGDLVEIRL------DFIEGF-RPREHLPLLL-
PhSDH3       LVCVPLVA---RTVEEMVADAAAAAAAGGDLVEIRL------DFIQGF-RPREHLPQLL-
SbSDH3       LLCVPLVG---RTVEEMIADAAAAAAATGGDMVEIRL------DFIQGF-RPREHLPPLF-
ZmSDH3       LLCVPLVG---RTVEEMITDAAAAAATGGDMVEIRL------DFIQGF-RPREHLPPLL-
TaSDH3       ------------------------------------------------------------
OsSDH2       LLCVPLVS---RTVEAMQADAAAAAAAAGADLVEIRL------DFIEGF-RPREHLPSLL-
BdSDH1       LLCVPLVA---RTVEEMETDAAAAAAAGGDLVEIRL------DFIEGF-RPREHLPRLL-
MaSDH3       ------------------------------------------------------------
MaSDH4       LLCVPLVA---KTVEQMTADMAAAKASGADLVELRV------DHLSAF-LPGRDLPLLL-
ElgSDH2      L---PPVA---KTVKQMAADMDAAKAAGADLVEIRL------AQPSAF-RPQEDLEFLL-
ElgSDH1      LLCVPLVA---KTVEQMAADMAAAKAAGADLVEIRV------DHLSTF-RPREDLEFLL-
PdSDH1       ------------------------------------------------------------
PdSDH2       LLCVPLVA---KTVEQMAADMAAAKAAGADVVEIRL------DHLSAF-RPRVDLEFLL-
MaSDH1       LLCVPLVS---KTVEQMVADMAAAKACGADLVEIRL------DHLSNF-DPRRDLQLLL-
MaSDH2       LLCVPLVA---KTVEQMMADMAAAKARGADLVEIRL------DHLSDF-EPRRDLPLLV-
PvSDH2       LICGSTTA---ESVEEMVFEMEKAKELGADLVEARL------DLLKDF-HPTHHLPFLI-
GmSDH1       LICASTSA---ESVEEMVFEMVKGKELGADLVEARL------DFLNDF-HPTQHLPSLI-
FvSDH2       LICAPVMG---ESVDQMLRQLQQAKELGADLVEIRL------DYIKNF-SPRQDLETLI-
```

```
PpSDH1       ------MA---ESVDQMLVQMGKAREVGADLVEIRL------DFLKNF-SPRQDLEILI-
PmSDH3       LICAPVMA---ESVDQMLVQMGKAREVGADLVEIRL------DFLKNF-SPRQDLEILI-
MdSDH1       LICAPVVA---DSVDQMLVQMGKAKEVGADLVELRV------DFLKNF-SPRQDLSVLI-
PbSDH1       LICAPVVA---DSVDQMLVQMGKAKEVGADLVELRV------DFLKNF-SPRQDLSVLI-
MdSDH4       ------------------------------------------------------------
PbSDH4       LICAPVMA---ESVDQMLVQMGKAKEVGADLVEIRV------DFLKHF-NPRQDLGTLI-
TcSDH3       LICVPIMA---ESVEEMLVKMRKAKELGGDLVEIRV------DFLKNF-IPKQDLDILI-
GrSDH3       LICVPIMA---ESVDQMLVQMQKAKDFGGDLVELRV------DFLKNF-VPRQDLDILI-
CasSDH3      MICAPVMA---ENVDQMLLMMRKAKELGADLVEVRI------DYLKNF-SPHHDLQILI-
DkSDH        MICAPVMA---ETAEQMLGQMRKAKELGADLVEIRI------DYLKSF-SPQQHLEVLI-
EgSDH2       LLCAPVMG---ESVDQMLGQIRAAKEQGADLVEVRL------DFLKSF-SPKQDLEILL-
VvSDH9       LICVPIMA---DSVDQMLGQIRKAKEVGGDLVEIRL------DYLKNF-SPRQDLQFLV-
SpSDH2       LICAPIMA---ESVDQMLIQMKRAKELGADIAEVRV------DFLKNF-SPRNDLETLI-
Poptr5       LICAPIMA---ESVDQMLVQMKRAKELGADVAEVRV------DFLKNF-SPRNDLEALI-
PeSDH3       LICAPIMA---ESVDQMLVQMKRAKELGADLAEVRV------DFLKNF-SPRNDLEALI-
MeSDH1       LICAPVMG---ESVDQMLVQMKKAKELGADLVEIRV------DFLKNF-SPRQNLEVLV-
JcSDH4       QICAPVMA---ESVDQMLVQMRKAKEYGADLVEIRL------DFLKNF-SPSQDLETLI-
CpSDH        LICVPIMA---ESVDKMVIDAGKAKASGADLVEIRL------DSLKSF-SPLDDLKTII-
AtSDH        LICAPVMA---DSIDKMVIETSKAHELGADLVEIRL------DWLKDF-NPLEDLKTII-
AlSDH        LICAPVMA---DSIDKMVIETFKAHELGADLVEIRL------DWLKDF-NPIEDLKTII-
PpSDH2       ------MA---ESVGEMVTDMGKAKGLGADVVEIRL------DYLKTF-NHNEDLKRLV-
PmSDH2       LVCVPIIA---ESVGEMVTDMGKAKGLGADLVEIRL------DYLKTF-NHNEDLKRLV-
PmSDH1       LVCVPIIA---ESVGEMVTDMGKAKGLGADLVEIRL------DYLKTF-NHNEDLKRLV-
MdSDH2       LLCVPIMA---ETVDQMVIDMGEAKAVGADLVEIRL------DXLKVF-NHNQDLKRLV-
PbSDH2       LLCVPIMA---ETVDQMVIDMGEAKAVGADLVEIRL------DCLKVF-NHNQDLKRLV-
MdSDH5       LICAPIMA------------------------------------------------------
FvSDH3       LVCAPIMA---ETVHKMVRDMSKARDLGADVVEIRL------DYLKVF-NSNQDLKTLI-
FsSDH        LVCAPIMA---ESVDKMVINMDKAKQGGADLVEIRL------DSLKSF-NPNGDLKTLI-
JrSDH        LVCAPIMA---ESVDKMVINMNKAKQGDADLVEIRL------DSLKSF-NPSNDLKTII-
FvSDH4       LICAPIMA---ESVAKMVVEMGRAKAVGADLVEIRL------DHLKVF-DSSEDVKTLI-
PpSDH3       LICAPIMA---ESVDKMLIDMDKAKTLGADLLEIRL------DHLKAF-NDNVDLKTLI-
MdSDH3       VICAPIMA---ESVDKMVIDMGKAKALGADLVEIRL------DHLKDF-NCNEDLKILI-
PbSDH3       VICAPIMA---ESVDKMVIDMGKAKALGADLVEVRL------DHLKDF-NCNEDLKILI-
NnSDH1       ------------------------------------------------------------
NnSDH2       FICAPIMG---QSVDQMLINMGKAKAAGADLVELRL------DYIKNF-NPYQDLEVMI-
SiSDH3       LICAPLMA---DSVDQVVVLMHTAKLKGADIVEIRL------DHLKSF-NGRADIEKLI-
NtSDH1       LICAPIMA---DTVDQMLNLMQKAKISGADLVEVRL------DSLKSF-NPQSDIDTII-
StSDH2       LICAPIMA---DSVDQMLILMQKAKTSGADLVEVRV------DSLKNF-NPRPDIDTLI-
SlSDH1       LICAPIMA---DSVDQMLILMQKAKISGADLVEVRV------DSLKSF-NPRPDIDTLI-
MtSDH        LICVPIMG---ETIQKMIADIQKAKLNGADLVEIRL------DSLKTF-NPSQDLNTFMQ
PvSDH3       LICVPIMG---ETIEKMKIDVQKAKAEGADVAEIRL------DSLSNI-DPYQDLTALI-
GmSDH3       LICVPIMG---ESVEKMEIDVDKAKAGGADLVEIRL------DSLKTF-DPYRDLNAFI-
LuSDH1       LICVPIMA---DSIDDMLVKASQAKTAGADLVEIRL------DSLTNFVHPRQDLTTLI-
LuSDH2       LICVPIMA---DSVDDMLVKASQAKTAGADLVEIRL------DSLTNFLHPRQDLTTLI-
TcSDH2       LICVPIMA---DSIAKMVVDMAKAKASSADLVEIRL------DTLNSF-NPYEDLKVLI-
GrSDH1       LICAPLMA---DSIDKMVTLMAKAKATTADLVEIRL------DSLKNF-NPFEDLNVLI-
SpSDH4       LICTPIMA---DSVDRMVILMAEAKSMGADVVEIRL------DSLKDF-NPNSDIKTLI-
Poptr1       LICTPIMA---DSVDKMAILMAEAKSVGADLVEIRL------DSLKDF-NPNSDIKTLI-
PeSDH1       LICTPIMA---DSVDKMAILMAEAKSVGADLVEIRL------DSLKDF-NPNSDIKTLI-
MeSDH2       MICAPIMA---DSVDKMIVYLDKAKATGADLAEIRL------DSLKSF-NPDEDLQILT-
RcSDH3       LICAPIMA---DSVDKMLINLVKAKENGADLVEIRL------DSLKTF-NASDDIKTLI-
JcSDH2       LICAPIMA---ESVDKMMIDFGKAKVSGADVVEIRL------DSLKTF-NPPDDLKTLI-
VvSDH5       LICVPIMG---ETIEKMVVDMSKAKTSGADLVEVRL------DTLKRF-NPRQDLEVLI-
EgSDH5       LICVPVMA---DSVEEMVIQLDKAKSSGADLVEIRV------DGLKNL-SPHEDLKTLI-
CsSDH3       LICVPIMG---ESVDKMVVDMGKANASGADLVEIRL------DGLKNF-NPRENIKTLI-
CcSDH1       LICVPIMG---ESVDKMVVDMGKANASGADLVEIRL------DGLKNF-NPRENIKTLI-
```

```
StSDH1       KQRSLPSIVSFR-------------------PNSAINS----D-WKKTCIQVLKLAVELD
SlSDH3       KQRSLPSIVSFR-------------------PKSPINS----E-GKKTCIQVLKLAVELD
SiSDH1       ------------------------------------------------------------
EgSDH1       RRRTLPAIVSFR-------------------LNSARASRRQ-N-DKTTCLQVLRLALELD
GrSDH4       ERKTLPAIVSFR-------------------LKSSRTLSKG-G-CYDTCLKVLKRALELD
CsSDH2       QHPTLPAIVSYR-------------------LKSSRKSSDE-A-CKNTCLQVLRRALDLD
CcSDH3       ------------------------------------------------------------
Poptr4       KQRTLPSIVSFR-------------------LEPSRISSNKDR--KNTCLQVLRLAFDLN
SpSDH3       KHRTLPSIVSFR-------------------LKPSGISSNRDR--KNTCLQLLRLAFDLN
JcSDH5       KLRTLPAIVSFR-------------------PKLSKISSNGDY-FNVTCRQVLRLAGNLD
RcSDH4       KLRTLPAIVSFR-------------------PKTSKISSNGDS--KKTCLQVLSLALDLD
MeSDH4       ----------------------------------------------------MRLAVDLN
SiSDH2       ENKPLPVVVVYG-------------------ATWEGGQYEGD---EQERLQALLLAKDLG
PvSDH1       QNKPLPLLIAHR-------------------PTWGGSLHKGD---EIMRLEALQLAVELG
GmSDH2       QNKPLPVLIVNR-------------------PKWEGGLYEGD---ENKRLEALQLAVELS
NtSDH2       NNNPLPVLIVYR-------------------PIWEGNEFEEDDDHIHKQLEVLRWAKELG
StSDH3       KNKPLPILIVYR-------------------PIWEGNDFEGD---AHKQLEVLGLAKELG
SlSDH2       KNKPLPILILYR-------------------PIWERNDFEAD---AHKQLEALRLAKELG
EgSDH4       KSKPLPVIIVYR-------------------PKCEGGLYEGD---ETARLEALHSALKLG
TcSDH1       KNKPLPVIIVYR-------------------PKWEGGQYEGD---ENSRLEALCLAREMG
GrSDH2       DNKPLPLIILYR-------------------PKWDGGLYEGD---EHSRLQALRLAAELG
CasSDH1      NNKPLPVLIVYR-------------------PQWEGGQYGGD---ENMRLDTLRLAKELG
VvSDH7       RNKPLPVMIVYR-------------------PKWEGGQYEGD---EHSRLEALHLAEKLG
CsSDH1       TKKPLPVLIVYR-------------------PKWAGGLYEGD---EHKRLEALHLAEDLG
CcSDH2       TKKPLPVLIVYR-------------------PKWAGGLYEGD---EHKRLEALHLAEDLG
SpSDH1       RNKPLPVIIVYR-------------------PKWEGGQYEGD---EHKRLEALRLAHDLG
Poptr3       RNKPLPVIIVYR-------------------PKWEGGQYEGD---EHRRLEALRLANDLG
PeSDH2       RNKPLPVIIVYR-------------------PKWEGGQYEGD---EHKRLEALRLAHDLG
Poptr2       RNKPLPVIIVYR-------------------PRWEGGQYEGD---EHTRLEALRLAHELG
SpSDH5       RNKPLPVIIVYR-------------------PRWEGGQYEGD---EHTRLEALRLAHDLG
RcSDH2       RNKPLPVIIVYR-------------------PKSEGGLYEGD---EPTRLEALRLAYVLG
JcSDH1       RKKPLPVIIVYR-------------------SKCEGGLFEGD---EHTRLEVLRLAHELG
SpSDH6       ------KLQSSP-------------------PVWEGGQY------------VLRLAMQFG
TcSDH4       KQSPLSTLATYR-------------------PIWEGGQYEGD---ETKRQDALRLAMELG
CasSDH2      KQSPLPTLVTYR-------------------PIWESGQYEGD---ENKRQDALRLAMELG
VvSDH8       KQSPLPTLVTYR-------------------PIWEGGQYEGD---ENKRQDALRLAMELG
FvSDH1       KQSPLPTLVTYR-------------------PVWEGGQYEGD---ETKRQDALRSAMELG
EgSDH3       KQSPLPTLVTYR-------------------PVWEGGQYEGD---ESKRQDALRLAMQLG
MeSDH3       KQSPLPTLVTYR-------------------PSWEGGQYEGD---ETKRQDALRLAMQLG
RcSDH1       KQSPLPTLVTYR-------------------PIWEGGQYEGD---ETKRQDALRFAMQLG
JcSDH3       KRSPLPTLVTYR-------------------PVWEGGQYEGD---ETKRQDALRLAMQLG
SbSDH2       KPRPLPALVTYR-------------------PKWEGGEYEGD---DESRFEALLLAMELG
ZmSDH2       KPRPLPALVTYR-------------------PKWEGGEYEGD---DESRFEALLLAMELG
SeiSDH2      EPRPLPALVTYR-------------------PKWEGGEYEGD---EEPRFEALMLAMELG
PavSDH2      EPRPLPALVTYR-------------------PKWEGGEYEGD---EEPRFEALMLAMELG
PhSDH2       EPRPLPALVTYR-------------------PKWEGGEYEGD---DERRLEALMLAMELG
PavSDH4      EPRPLPALVTYR-------------------PKWEGGEYEGD---EEPRFEALMLAMELG
OsSDH1       QPRPLPALVTYR-------------------PKWEGGEYEGD---DEPRFEALLLAMEMG
BdSDH2       KPRALPALVTYR-------------------PKWEGGEYDGD---DEPRFEALQLAMELG
TaSDH1       ------------------------------------------------------MELG
TaSDH2       --MYVAYCWFHR-------------------PKWEGGEYEGD---DEPRFEALQLAMELG
ZmSDH1       QPRSLPVIVTYR-------------------PNWEGGQYAGE---DEPRFEALLLAMELG
SbSDH1       QPRSLPVIVTYR-------------------PKWEGGQYAGE---DEPRFEELLLAMELG
SeiSDH1      QPRPLPVIITYR-------------------PKWEGGEYEGE---DEPRFEALLLAMELG
PavSDH3      QPRPLPVIVTYR-------------------PKWEGGQYDGE---DQPRFEALLLAMELG
PhSDH1       QPRPLPVIVAYR-------------------PKWEGGQYEGE---DETRFETLLLAMELG
PavSDH1      QPRPLPVIVTYR-------------------PKWEGGQYEGE---DQPRFEALLLAMELG
SeiSDH3      RGCPLPALVTYR-------------------PNWEGGQYEGD---DTTRFEALHIAMELG
PhSDH3       RGCPLPALVTYR-------------------PTWEGGQYEGD---DATRFETLRLAMELG
SbSDH3       RSCPLPALVTYR-------------------PVWEGGQYEGD---DTTRFETLRLAMELG
ZmSDH3       RGCPLPALVTYR-------------------PVWEGGQYEGD---DTTRFETLRLAMELG
TaSDH3       ------------------------------------------------------MELG
OsSDH2       RGCPLPALVTYR-------------------PNWEGGQYDGD---DATRFEALRLAMELG
BdSDH1       HGCPLPALVTYR-------------------PNWEGGRYDGD---DATRFETLRLAMELG
MaSDH3       ------------------------------------------------------MELG
MaSDH4       RDRPLPALVTYR-------------------PKWEGGEYEGD---DKQRLEALLLAMELG
ElgSDH2      KGRPLPALVTYR-------------------PKWEGGEYEGD---DNQRFEALHTAVDLG
ElgSDH1      KGRPLPAIVTYR-------------------PKWEGGEYEGD---DNRRFEALRIAMDLG
PdSDH1       ------------------------------------------------------MDLG
PdSDH2       KGRPLPALVTYR-------------------PKWEGGEYEGD---DNQRFEALRFAMDLG
MaSDH1       DDRPLPALVTYR-------------------PKWEGGEYEGN---DKQRFQALCLAMELG
MaSDH2       GDRPLPVLVTYR-------------------PKWEGGEYEGD---DKQRFEALCLAMELG
PvSDH2       KNRHLPIIISYR-------------------PIWEGGEYKGD---ETMRQEALRLAIHLE
GmSDH1       NNRPLPILITYR-------------------PIWEGGEYDGD---ESQRQDALRLAIELG
FvSDH2       KRSPLPTLVTYR-------------------PKWEGGQYEGD---EKKRQEALILAMELG
```

```
PpSDH1       KHSPLPTLFTYR--------------------PAWEGGQYKGD---DNKRQDALHLAMELG
PmSDH3       KHSPLPTLFTYR--------------------PAWEGGQYKGD---DNKRQDALRLAMELG
MdSDH1       KQSPLPTLVTYR--------------------PAWEGGQYKGD---DKQRQDALRVAMDLG
PbSDH1       KQSPLPTLVTYR--------------------PAWEGGQYKGD---DKQRQDALRVAMDLG
MdSDH4       ------------------------------------------------------------MELG
PbSDH4       NQSPLPTLVTYR--------------------PSWEGGQYKGD---DKQRQDALRVAMELG
TcSDH3       KQAPLPTLVTFR--------------------PRWEGGRYDGD---ESKRQEALRLAMELG
GrSDH3       KQAPLPTLVTYR--------------------PRWEGGQYDGD---ESKRREALRLAMELG
CasSDH3      KQCPLPTLITYR--------------------PTWEGGQFDGD---ETRRQAALRQAVELG
DkSDH        KQSPLPTIITYR--------------------PTWEGGQYDGD---ESRRQATLHQAMELG
EgSDH2       KQSALPTLVTYR--------------------PKWEGGQYEGD---DSRRLDALRLALELG
VvSDH9       KQSPLPTLVTYR--------------------PTWEGGQYDGD---EGKRLDALRLAIELG
SpSDH2       KQGPLPTLITYR--------------------PKWEGGQYDGD---ENKRQKALQLAMELG
Poptr5       KQCPLPTLITYR--------------------PKWEGGQYDGD---ENKRQKALQIAMELG
PeSDH3       KQCPLPTLITYR--------------------PKWEGGQYDGD---EDKRQKALQIAMELG
MeSDH1       KQSPLPTLITYR--------------------PKWEGGQYDGN---ENKRQEALRLAMELG
JcSDH4       KQSPLPTLITYR--------------------PKWEGGQYDGD---ESKRQEAFRLAMELG
CpSDH        TECPLPTLFTYR--------------------PKWEGGQYDGD---EKERLNTLRLAMELG
AtSDH        KKSPLPTLFTYR--------------------PKWEGGQYEGD---ENERRDVLRLAMELG
AlSDH        KKSPLPTLFTYR--------------------PKWEGGQYEGD---ENERQDVLRLAMELG
PpSDH2       KECPLPTLFSYR--------------------PKWEGGQYDGD---EQNRLDALRLAMDLG
PmSDH2       KECPLPTLFSYR--------------------PKWEGGQYDGD---EQNRLDALRLAMDLG
PmSDH1       KECPLPTLFSYR--------------------PKWEGGQYDGD---EQNRLDALRLAMDLG
MdSDH2       KECPLPTLFSYR--------------------PIWEGGQYDGD---EKTRLDALRLAMELG
PbSDH2       KECPLPTLFSYR--------------------PIWEGGQYDGD---EKTRLDALRLAIELG
MdSDH5       -----------E--------------------PKWEGGQYDGD---EKYRLDALRLAMELG
FvSDH3       KESPMPTLFTYR--------------------PKWEGGQYDGD---EKHRQDILRLAMELG
FsSDH        KECPLPTLFTYR--------------------PKWEGGQYDGD---EKKRLDALRLAMEFG
JrSDH        KASPLPTLFTYR--------------------PKWEGGQYDGD---EKKRLDALRLAMEFG
FvSDH4       DQSPLPTLFTYR--------------------PKWEGGQYDGD---EKSRLDALRLAMELG
PpSDH3       KESPLPTLFTYR--------------------PKWEGGQYDGD---EKDRLDALRLAMELG
MdSDH3       KESPLPTLVTYR--------------------PKWEGGQYDGD---EKYRLDALRLAMEXG
PbSDH3       KESPLPTLITYR--------------------PKWEGGQYDGD---EKYRLDALRLAMELG
NnSDH1       ------------------------------------------------------------
NnSDH2       KQCPLPTLCTYR--------------------PKWEGGEFEGD---EKRRTDALRLAMELG
SiSDH3       RECPLPTLFTYR--------------------PIWEGGQYDGD---ENSRFNALRLAMELG
NtSDH1       KQSPLPTLFTYR--------------------PTWEGGQYAGD---EVSRLDALRVAMELG
StSDH2       KQCPLPTLFTYR--------------------PIWEGGQYAGD---EMSRLDALRLAMELG
SlSDH1       KQCPLPTLFTYSYVLGVGQGILLIRYYKGIGPTWEGGQYAGD---EKSRLDALRLAMELG
MtSDH        QHHSLPFLFTYR--------------------PKWEGGMYDGD---ENQRLDALQLAMELE
PvSDH3       QHRVLPLLFTYR--------------------PKWEGGMYDGD---ENKRLDALRLAMELG
GmSDH3       QHRSLPLLFTYR--------------------PKWEGGMYDGD---ENKRLDALRLAMELG
LuSDH1       NGSPLPTLCTYR--------------------PRWEGGQYDGD---ESERLDVLKLAMELG
LuSDH2       NGSPLPTLCTYR--------------------PRWEGGQYGGD---ESERLDVLKLAMELG
TcSDH2       KECPLPTLFTYR--------------------PVWEGGQYDGD---EKERLDVLRLAMELG
GrSDH1       KQSPLPTLFTYR--------------------PVWEGGQYDGD---EKKRLDVLRLAMELG
SpSDH4       LHSPLPTLFTYR--------------------PMWEGGQYNGD---EKPRLDALRLAMELG
Poptr1       LHSPLPTLFTYR--------------------PMWEGGQYNGD---EKPRLDALRLAMELG
PeSDH1       LHSPLPTLFTYR--------------------PMWEGGQYNGD---EKPRLDALRLAMELG
MeSDH2       TRSPLPTLFTYR--------------------PIWEGGQYEGD---ETSRLDALRLAMELG
RcSDH3       TRSPLPTLFTYR--------------------PKWEGGEYDGD---ENERLDALRLAMELG
JcSDH2       TRSPLPTIFTYR--------------------PKWEGGQYEGD---ENKRLDVLRSAMELG
VvSDH5       RKCPLPTLFTYR--------------------PKWEGGQYEGD---ENSRRDALRLAMELG
EgSDH5       KASALPTLFTYR--------------------PKWEGGQYDGD---DKPRLETLRLAMELG
CsSDH3       KESPVPTLFTYR--------------------PIWEGGQYDGD---ENERVDVLRLAMELG
CcSDH1       KESPVPTLFTYR--------------------PIWEGGQYDGD---ENERVDVLRLAMELG
```

```
StSDH1        VEFVEVDTE-VVSHQV--VAELMSSRSNSKIIASTYVINAGNPPTKDALCNSIINLQSTG
SlSDH3        VEFVEVDTQ-VVCHQV--VAELMKSRSNSKIIASTYVN--GGNPTKDTLCNSIINLQSTG
SiSDH1        -------MLQVCTKEL--INELKRQRSKTRIIVSNHVNK-G-GFTREKLADLIIGMQSTD
EgSDH1        VEFVEMEHE-VVSHFN--IDELMEKRSSSKIIVSRHLNG--DKPCKEKLGNLIALMQSSG
GrSDH4        VEFVEMDYE-VACEVN--MAEYLHNRSNTRIIVSNYVND--GRPSVEKLVDIIACMQATG
CsSDH2        VEFVEMDYE-VASDPL--MSEIIYSRSNTKIIVSSYLNG-GGKPTTEKLGDVIACMQATG
CcSDH3        -----MDYE-VASDPL--MSEIIYSRSNTKIIVSSYLNG-GGKPTTEKLGDVIACMQATG
Poptr4        VEFVEMDYE-VASEDV--MAEYVYNRSNTKLIVSSYVNG--RKPSAEELGYLIACMQSTG
SpSDH3        VEFVEMDYE-VASEDV--MAEYVYNRSSTKLIVSSYVNG--TKPSAEELGYLIACMQSTG
JcSDH5        VEFVEIDYPQVVSDID--MAEYVYNRSNTKLIVSSYVKG--EKPSAEKLGNLIVHMQSTG
RcSDH4        VEFVEMEYE-VASDVS--MAEYIYNRSSGKLIVSSYVNG--GKPSAEKLGNLIANLQSTG
MeSDH4        VEFVEMDYE-VASDTN--MAECVYNRPNSKLIVSSYVNG--GKPSPEKLGNLIACMQSTG
SiSDH2        ADYIELELK-VASAFM--KEDKLGRLKSSKIIVSCYLDS--PASLREDLGRLVAEMQSMN
PvSDH1        ADFIEVELK-EASCLPTLVEHRRNHNSHGKIIVSCYVDD--ITPTQEELLQLVALMQATG
GmSDH2        ADFIDVELK-AASCLPTLVEHMRNHNSHGKIIVSCYVDG--TTPPHEVLLQLVELMQATG
NtSDH2        ADYIELDLK-IASDFT--KKEKPRWSSGCKVIASCFVDN--VT-SKEDLSQVVAHMQSTG
StSDH3        ADYVELDLK-IASDFA--KNEKSSWSSGCKLITSCFVDN--VT-SKEDLSQIVASMQSTG
SlSDH2        ADYVELDLK-IASEFA--KNEKSSWSSGCKLITSCFVDN--VT-SKEDLSQVVASMQSTG
EgSDH4        ADYVDFELK-VAAELM--SKQNKLHCGGTKVIVSCFLDG--VTPSKEELSNLATHMQATG
TcSDH1        ADYIDFELK-VASDLI--REQKMKNDNATKVIVSQNIDG--MTPKDEELSNLAASIQATG
GrSDH2        ADYIDFELK-VASHLI--RELKINYQSATKFIVSCHITG--TTPSEEELANLAATMRATG
CasSDH1       ADYIDLELK-VASHLM--EEHNSNKHRSSKIIVSRYIDG--TTPSEEDLSQLVAHMQSTG
VvSDH7        ADYIDFELK-VASDFL--GKQKMDQHSSSRTIVSCYVDG--VTPPIEDLICRVALLQSTG
CsSDH1        ADYVDFELK-VASNIL--GKQYSSHQSGTRFIVSCNLDC--ETPSEEDLGYLVSRMQATG
CcSDH2        ADYVDFELK-VASNIL--GKQYSSHQSGTRFIVSCNLDC--ETPSEEDLGYLVSRMQATG
SpSDH1        ADYIDFELK-VASELI--GELKNKHRNGGKVIVSSYLNG--ATPSKEKLGDLVASMQATG
Poptr3        ADYIDLELK-VASELI--WELKNKHQNGGKVIVSSYLNG--ATPSKENLSHLVATMQATE
PeSDH2        ADYIDLELK-VASELI--WELKNKHQNGGKVIVSSYLNG--ATPSKENLSHLVATMQATE
Poptr2        ADYIDVELK-VASDLV--REVKNKHQTGGKVIVSSYLSG--ATPSKEDLSHLVASMQATK
SpSDH5        ADYIDVELK-VASDLV--GEVKNKHQAGGKVIVSSYLNG--ATPSKEDLGHLVARMQSTG
RcSDH2        ADYVDFELK-VASDLI--GELKGTHHTGSKVIVSCYVNG--DMPTKENLSQLVASMQATG
JcSDH1        ADYIDLELK-VASDMI--GELKKTHQTDSKIIVSCYLTG--ATPPKENLSHLVGSMQAAG
SpSDH6        ANYVDVELE-------------------------------------AIASLVARIQATG
TcSDH4        ANYIDIELE-VAHDFI--SSIYGKKPDNIKVIVSSHNFH--NTPSTEAIGNLVARIQATG
CasSDH2       ADYIDVELQ-VAQEFN--NSISAKKPEKFKVIVSSQNFH--STPSAEAIGNLVARIQATG
VvSDH8        ASYIDVELE-VAHEFN--NSIYGKKPQNFKVIVSSHNFH--NTPSTEAIGNLVARIQASG
FvSDH1        ANYIDIELE-VAHEFN--NSIYEKKPDNFKVIVSSHNFH--NTPSSEAIGNLVARIQATG
EgSDH3        AHYIDVELE-VAHDFN--SSILGKKPDNFKIIVSSHNFH--NTPSAEAIGNLVARIQATG
MeSDH3        ATYVDIELQ-VAHDFN--NSISGSKPDNFKVIVSSHNFH--NTPSAEAIANLVARIQATG
RcSDH1        ANYVDVELE-VAHEFN--NSIYGKKPDNFKVIVSSHNFH--NTPSSEAIANLVARIQATG
JcSDH3        ANYVDVELE-VAHDFN--SSIYGKKPDNFRVIVSSHNFH--NTPSAEAIANLVARIQATG
SbSDH2        AEYVDVEFK-VADKFM--KLISGKKPDNCKLIVSSHNYE--NTPSSEELANLVAQIQATG
ZmSDH2        AEYVDVELK-VADKFM--KLISGRKPDNCKLIVSSHNYE--TTPSSEELANLVAQIQATG
SeiSDH2       AEYVDIELK-VADKFM--RLLSGKKPENCKLIVSSHNYE--NTPSAEELANLVAQIKATG
PavSDH2       AEYVDIELK-AAEKFM--RLLAGKKPENCKLIVSSHNYE--NTPSAEELADLVAQIQATG
PhSDH2        AEYVDIELK-AADKFM--RLLSGKKPENCKLIVSSHNYE--NTPSAEELADLVAQIQATG
PavSDH4       AEYVDMELK-AANKFM--RLLAGKKPENCKLIVSSHNYE--NTPSVEELADLVAQIQATG
OsSDH1        AEYVDIELK-VADKFM--KLISGKKPETCKLIVSSHNFE--NTPSTEELENLVAQIQATG
BdSDH2        AEYVDIELK-VADKFM--KFMSGKKPEKCKLIVSSHNYD--NTPSAEELGSLLAQIQATG
TaSDH1        AEYVDIELK-VADKFI--KFMSGKKPENCKLIVSSHNYD--NTPSAEELASLLAQIQATG
TaSDH2        AEYVDIELK-VADKFI--KFMSGKKPENCKLIVSSHNYD--NTPSAEELASLLAQIQATG
ZmSDH1        AEYVDVEFK-VADKFM--KYLSGKKPQNCKLIISAYDYK--TTPSVHELLNLVAQIRATG
SbSDH1        AEYVDIEFK-VADKFM--KFLSGKKPQSCKLIVSAYDYE--RTPSVNELLNLVAQIRAIG
SeiSDH1       AEYVDIEHK-VVVNFL--KFLSGRKPETCKLIVSIHNYE--YTPSVDELLSLVDQIQATG
PavSDH3       AEYVDIEFK-VADKFM--NFLSGRKPETCKLIVSFHNYD--YTPSVEELLSLVDQIQATG
PhSDH1        AEYVDIEFK-VADKFL--KFLSGRKPETCKLIVSFHNYE--CTPSGEELLSLVDQIQATG
PavSDH1       AEYVDIEFK-VADKFL--NFLSGKKPETCKLIVSFHNYE--YTPSVEELLSLVDQIQATG
SeiSDH3       VDYVDVELK-VADKFM--NFISGNKPEKCKLIVSSHNYE--YTPSCEELANLVASIQAVG
PhSDH3        VDYVDVELK-VADKFM--SFISGNKPEKCKLIVSSHNYE--YTPSCEELANLVAVIQAVG
SbSDH3        VDYVDIELK-VADKFM--SFISGHKPEKCKLIVSSHNYE--YTPSCEEITNLVARIQAVG
ZmSDH3        VDYVDIELK-VADKFM--SFIAGHKPEKCKLIVSSHNYE--YTPSCEEITNLVASIQAVG
TaSDH3        VDYVDIELK-VADKFV--DFLSGNKPDKCKLIVSSHNYE--STPSCEELANLVARIQEVG
OsSDH2        VDYVDIELK-VADKFI--SFISGSKPEKCKLIVSSHNYE--STPSCEELADLVARIQAVG
BdSDH1        AVYVDIELK-VADKFI--NFLSGYKPEKCKLIVSSHNYE--STPSCEELANLVARIQAVG
MaSDH3        ADYVDVELK-VAEDFV--RLIDGKKPENFKLIVSSHNYQ--KTPSSEELGSLVAAIQASG
MaSDH4        ADYVDVELK-VAEDFV--RLIDGKKPENFKLIVSSHNYQ--KTPSSEELGSLVAAIQASG
ElgSDH2       AEHVDIELK-VAEEFV--RLISGKKPENFKLIVSSHNYQ--STPSSEELGSLAARTQAVG
ElgSDH1       AEYVDIELK-VAEEFV--RLISGKKPENFKLIVSSHNYQ--STPSSEELGSLVARIQAVG
PdSDH1        AEYVDIELK-VAEEFV--RLISGKKPEDFKLIVSSHNYQ--STPSSEELGSLVARIQAVG
PdSDH2        AEYVDIELK-VAEEFV--RLISGKKPENFKLIVSSHNYQ--STPSSEELGSLVARIQAVG
MaSDH1        AEYVDVELK-VADDFV--RLISEKKPENFKLIISSHNYQ--STPSNEELSSLVARIQAAG
MaSDH2        AEYVDVELK-VADDFV--RLISGKKPDNFKLIVSSHNYQ--STPSSEELSSLVARIQAAG
PvSDH2        CEFVEVELK-VADEFY--NRIGWKKAESVKIIVSSHNLE--NTPSMEEIGNLAATIQASG
GmSDH1        SEFVDVELK-VADEFY--KSIGGKKAESVKIIVSSHNLE--STPSVEEIGNLAARIQATG
FvSDH2        ADYIDVELK-VANDFY--SSIQGKKPERVKIIVSSHNYE--STPSAEEIGNLVATIQATG
```

```
PpSDH1       ADYIDVELQ-VAHEFY--NSIQGEKPEKVKIIVSSHNYQ--KTPSTEEIGDLVARIQATG
PmSDH3       ADYIDVELQ-VAHDFY--NSIQGEKPEKVKIIVSSHNYQ--KTPSTEEIGDLVARIQATG
MdSDH1       ADFIDVELK-VADEFY--NSIQGRKPEKVKIIVSSHNYE--KTPSAEEXGDLVARIQATG
PbSDH1       ADFIDVELK-VADEFY--NSIQGRKPEKVKIIVSSHNYE--KTPSGEEIGDLVARIQATG
MdSDH4       ADFIDVELK-VADEFY--NSIKGKKPEKVKIIVSSHNYE--KTPSSEEIGDLVARIQATG
PbSDH4       ADFIDVELK-VADEFY--NSIQGKKPEKVKIIVSSHNYE--KTPSSEEIGDLVARIQATG
TcSDH3       ADYIDIELK-VAHDFF--SSLPRDRPENVKIIVSSHNYE--STPSVEELGDLVARIQATG
GrSDH3       ADYIDIELK-VAHDFF--NSLPGKRPEKVKIIVSSHNYE--RTPSVKELSHLVARIQATG
CasSDH3      ADYIDIELK-VADEFY--KLIQGKKPEKVKIIVSSHNYE--NTPSAEEIGNLAARIQATG
DkSDH        ADYIDIELK-VADEFF--SSIQERSPKRPKVIVSSHNYE--NTPSAEEIGNIAARIQATG
EgSDH2       ADYVDVELQ-VAQEFF--SSIQGKKPEKAKIIVSSHNYQ--NTPSSEELGNLVAKIQATG
VvSDH9       ADYIDVELQ-VAPEFI--NSIQGKTSGKVKIIVSSHNYQ--NTPSAEELGNLVARIQATG
SpSDH2       ADFIDIELK-VAQEFY--NFIQGKKPEKVKIIVSSHNYE--CTPSIEDIGELAVRIQATG
Poptr5       ADFIDIELK-VAQEFY--NFIQGKKPEKVKIIVSSHNYE--CTPSIEEIGDLVARIQATG
PeSDH3       ADFIDIELK-VAQEFY--NFIQGKKPEKVKIIVSSHNYE--CTPSIEEIGDLVARIQAAG
MeSDH1       SDFIDVELK-VAQEFF--SSIQGKKPEKVKIIVSSHNYD--RTPTVEEIGELVARIQATG
JcSDH4       SDFIDVELK-AAQEFF--SSIQGKKPEKSKIIVSSHNYE--CTPSVEDIADLVARIQATG
CpSDH        ADYVDVELQ-VAGEFV--NSIYEKKPEKFKVIVSSHNYQ--NTPSAEDLSNLVAKIQSTG
AtSDH        ADYIDVELQ-VASEFI--KSIDGKKPGKFKVIVSSHNYQ--NTPSVEDLDGLVARIQQTG
AlSDH        ADYVDVELQ-VASDFI--KSIDGKKPESFRVIVSSHNYQ--NTPSVEDLNDLVARIQQTG
PpSDH2       ADYVDVELQ-VAREFI--SSIEGKKPERLKVIVSSYNFE--ETPSVESIGNLVAAMQATG
PmSDH2       ADYVDVELQ-VAREFI--SSIQGKKPEKLKVIVSSYNFE--ETPSVESIGNLVAAMQATG
PmSDH1       ADYVDVELQ-VAREFI--SSIQGKKPEKLKVIVSSYNFE--ETPSVESIGNLVAAMQATG
MdSDH2       ADYIDVELQ-VAREFI--DSINGKRPETLKVIVSSHNFQ--DTPSVEALGNLVARIQATG
PbSDH2       ADYIDVELQ-VAREFI--DSINGKKPETLKVIVSSHNFQ--DTPSVEALGNLVARIQASG
MdSDH5       ADYIDVELQ-VAHEFI--DSIYGKKPEKFKVIVSSHNYQ--ETPSVEALGNLVARIQATG
FvSDH3       ADYIDVEFQ-VAHEFI--DSINGNKPEKLKVIVSSHNYQ--DTPSVKDLGDLVARIQATG
FsSDH        ADYIDVELQ-VAQEFN--ESIYGKKPEMFKVIVSSHNYQ--NTPSVEDLGNLVARIQATG
JrSDH        ADYIDVELQ-VACEFN--DSIYGRKPENSKVIVSSHNYQ--DTPSAEDLGNLVARIQATG
FvSDH4       ADYIDVELQ-VAQEFV--DFIRDKKPEKFKVIVSSHNYQ--ETPSVEALGNLVAAIQATG
PpSDH3       ADYIDVELQ-VAHEFI--DSIYGKKPEKFKVIVSSHNYQ--HTPSVEALGNLVARIQATG
MdSDH3       ADFIDVELQ-VAHEFI--DSIYGKKPEKFKVIVSSHNYQ--ETPSVEALGNLVAKIQATG
PbSDH3       ADFIDVELQ-VAHEFI--DSIYGKKPEKFKVIVSSHNYQ--ETPSVEALGNLVAKIQATG
NnSDH1       ------------------------------------------------------------
NnSDH2       ADYIDIELQ-VAHEFI--NSICGRKPQNLKIIVSSHNYQ--NTPSVEGIGDLVARIQATG
SiSDH3       ADHIDIELK-AAHEFN--NFIRGNKPEKCKVIVSSHNYD--NTPSAEDLGNLVARIQAAG
NtSDH1       ADYIDVELK-AIDEFN--TALHGNKSAKCKVIVSSHNYD--NTPSSEELGNLVARIQASG
StSDH2       ADYIDVELK-AIDEFN--NALHGNKSAKCKLIVSSHNYE--STPSAEDLGNLVARIQASG
SlSDH1       ADYIDVELK-AIGEFN--NALHGNKSAKCKLIVSSHNYE--STPSAEDLGNLVARIQASG
MtSDH        ADYIDVELK-VAHQFY--DSIRGKTYNKTKVIVSSHNYQ--QTPSVEDLGNLVARIQATG
PvSDH3       ADYIDVELQ-VAHAFY--DSIRGKTFNKTKVIVSSHNYQ--LTPSIEDLGNLVARIQATG
GmSDH3       ADYIDIELQ-VAHEFY--DSIRGKTFNKTKVIVSSHNYQ--LTPSIEDLGNLVARIQATG
LuSDH1       ADYIDVELQ-VAGDFI--ESLHGNRPPKCKVIVSSHNYE--NTPSVEDLGNLVARIQSAG
LuSDH2       ADYIDVELQ-VAGDFI--ESLHGNRPAKCKVIVSSHNYE--NTPSVEDLGNLVARIQSTG
TcSDH2       ADYIDVELK-VAHDFM--KSIDGKKPEKLKVIVSSHNYQ--STPSVEELGNLVARIQSTR
GrSDH1       ADYIDVELK-VAHEFI--KSIGGKKPEKIKVIVSSHNYQ--STPSVVELGNLVVKIQSTG
SpSDH4       ADYIDVELK-VALEFN--ELLRGNKPGKCKLIVSSHNYE--NTPSVEELGNLVARIQAAG
Poptr1       ADYIDVELK-VAHEFN--ELLRGNKPGKCKLIVSSHNYE--NTPSVEELGNLVARIQAAG
PeSDH1       ADYIDVELK-VAHEFN--DLLRGNKPGKCKLIVSSHNYE--NTPSVEELGNLVARIQAAG
MeSDH2       ADYIDVELQ-VARQFN--DSIRGRKPAKCKVIVSSHNYQ--STASVEELGNLVAKIQAAG
RcSDH3       ADYIDVELK-VAHKFN--DSIRGNKPAKCKVIVSSHNYH--NTPSVEDLGNLVADIQAAG
JcSDH2       ADYVDVELQ-VAREFN--DSILGSKPAKCKIIVSNHNYH--NTPSVEELGNLVARIQAAG
VvSDH5       ADYVDIELK-VAHEFI--NSIHGRKPEKFKVIVSSHNYQ--NTPSVEDLGNLVVSIQATG
EgSDH5       ADYIDVELK-VASEFN--ASIQGRKPEKCKVIVSSHNYE--NTPSAEDLSNLVARIQAAG
CsSDH3       ADYIDVELQ-VAREFN--DSIRGKKPEKCKVIVSSHNYQ--YTPSVEDLSNLVARIQASG
CcSDH1       ADYIDVELQ-VAREFN--DSIRGKKPEKCKVIVSSHNYQ--YTPSMEDLSNLVARIQASG
```

```
StSDH1       ADIIKLVIDVAYITDVAPVFHMLTHSQ---------------------------------
SlSDH3       ADIIKLVIDVAYITDVAPVFHMLTHSQ---------------------------------
SiSDH1       ADIIKLVINVESITDVAPLFHMLTHCQ---------------------------------
EgSDH1       GDVIKLLVDVDYITDLAPIFQLLTSCQ---------------------------------
GrSDH4       ADVIKLDLCVDYITDLAPIFTVLTHCQ---------------------------------
CsSDH2       ADVMKLEIAVDSITDLAPVFEMLTHCQ---------------------------------
CcSDH3       ADVMKLEIAVDSITDLAPVFEMLTHCQ---------------------------------
Poptr4       ADVLKLVLDVEKITDLAPVFTMLTHCQ---------------------------------
SpSDH3       ADVLKLVLDVEKITDLAPVFTMLSHCQ---------------------------------
JcSDH5       ADVIKLVINVDYITDLAPIFKMLAHCQ---------------------------------
RcSDH4       ADVIKLVINVDYITDLAPVFKMLAHCQ---------------------------------
MeSDH4       ADVIKLVINVEYITDLASVFKILAHCQ---------------------------------
SiSDH2       ADIIKFVSGAANITEIMEVFELLSHSQ---------------------------------
PvSDH1       SDIIKLVTHAADITEVMRIFSLFPHCQ---------------------------------
GmSDH2       ADIIKLVTLAADITEIKRIFSLFLYCQ---------------------------------
NtSDH2       ADILKIVTNANDITELEKMFHLLSHCQ---------------------------------
StSDH3       ADILKIVTNANDITELEKTFHLLSHCQ---------------------------------
SlSDH2       ADILKIVINANDITELEKTFHLLSHCQ---------------------------------
EgSDH4       ADIIKIVTSASNITELARLFHILSYSQ---------------------------------
TcSDH1       ADIIKVVVNVADITEIAKIFHLLSHYQ---------------------------------
GrSDH2       ADILKVVVNVTDITEIARIFHLLSHCQ---------------------------------
CasSDH1      ADIIKLVSKSSSITELPRFFHLLSHCQ---------------------------------
VvSDH7       ADMIKLVINATNITEITKIFHLLSHCQ---------------------------------
CsSDH1       ADIIKLVFSVNDITEIARIFQLLSHCQ---------------------------------
CcSDH2       ADIIKLVFSVNDITEIARIFQLLSHCQ---------------------------------
SpSDH1       ADIIKVVSNADDITEMERIFYLLSHCE---------------------------------
Poptr3       ADIIKVVSNADDITEMERIFHLLSHCE---------------------------------
PeSDH2       ADIIKVVSNADDITEMERIFYLLSHCE---------------------------------
Poptr2       ADIIKVVSNANDITELDRIFHLLSHMQ---------------------------------
SpSDH5       ADIIKVVSNADDITELERIFHLLSRSE---------------------------------
RcSDH2       ADIIKLVSTANNITELDRIFHLILHCQ---------------------------------
JcSDH1       ADIIKLVSNANNITELQRIFYLLLHSQ---------------------------------
SpSDH6       ADIVKIATAALDVTDCARIFQIMVHSQ---------------------------------
TcSDH4       ADIAKIAATALDITDCARICQIT-------------------------------------
CasSDH2      ADIVKIVTTALDITDVARIFQITVHSQ---------------------------------
VvSDH8       ADIVKIATTALDITDVARVLQVTVHSQ---------------------------------
FvSDH1       ADIVKIATTALDITDCARIFQITVHSQ---------------------------------
EgSDH3       ADIAKIATTALDITDCARIFQITVHSQ---------------------------------
MeSDH3       ADIVKIATTALDITDCARIFQIMVHSQ---------------------------------
RcSDH1       ADIVKIATTALDITDCARIFQIMVHCQ---------------------------------
JcSDH3       ADIIKIATTALDITDCARIFRIMVHSQ---------------------------------
SbSDH2       ADIVKIATTATEIVDVAKMFQILVHCQE--------------------------------
ZmSDH2       ADIVKIATTATEIVDVAKMFQILVHCQE--------------------------------
SeiSDH2      ADIVKIATTATDIVDVARMFQILVHCQE--------------------------------
PavSDH2      ADIVKIATTATEIVDVARMFQILVHCQE--------------------------------
PhSDH2       ADIVKIATTATEIVDVARMFQILAHCQE--------------------------------
PavSDH4      ADIVKIATTATEIVDVARMFQILVHCQE--------------------------------
OsSDH1       ADIVKIATTATEIVDVAKMFQILVHCQE--------------------------------
BdSDH2       ADIVKIATTATEIVDVSRMFQILVHCQE--------------------------------
TaSDH1       ADIVKIATTATEIVDVSRMFQILVHCQE--------------------------------
TaSDH2       ADIVKIATTATEIVDVSRMFQILVHCQE--------------------------------
ZmSDH1       ADIVKIETTAADIVDTSRMFQVLAHCHE--------------------------------
SbSDH1       ADIVKISTTAADIVDASRMFQVLAHCQE--------------------------------
SeiSDH1      ADIVKIATSATEIDDVSRMFQVLVHCQA--------------------------------
PavSDH3      ADIVKIATTATEIDDVSRMFQVLVHCKA--------------------------------
PhSDH1       ADIVKIATAATEIDDVSRMFQVLVHCKV--------------------------------
PavSDH1      ADIVKIATTATEIDDVSRMFQVLVHCKA--------------------------------
SeiSDH3      ADIVKIATTAKDIVDVSRMFQVMVHCQ---------------------------------
PhSDH3       ADIVKVATTAKDIVDVSRMFQVMVHCQ---------------------------------
SbSDH3       ADIVKVATTAKDIVDVSRMFQVMVHCQ---------------------------------
ZmSDH3       ADIVKVATTAKDIVDVSRMFQVMVHCQ---------------------------------
TaSDH3       ADIVKVATTATDIIDVSRMFQVMVHCQ---------------------------------
OsSDH2       SDIVKIATTASDIADVSRMFQVMVHCQ---------------------------------
BdSDH1       ADIVKVATTATDIVDVSRMFQVMVHCQ---------------------------------
MaSDH3       ADIVKIATTAVSVVDVTRMFQVLVHCQ---------------------------------
MaSDH4       ADIVKIATTAVSVVDVTRMFQVLVHCQ---------------------------------
ElgSDH2      ADIVKVVTTAVDIVDVTRMFQLIVHCQE--------------------------------
ElgSDH1      ADIVKIVTTAVDIVDVTRMLQVIVHCQ---------------------------------
PdSDH1       ADIVKIATTAVDIVDVTRMFQVIVLCQ---------------------------------
PdSDH2       ADIVKIATTAVDIVDVTRMFQVIVHCQE--------------------------------
MaSDH1       ADIVKIATTAVDIVDVARMFQVIVHCQ---------------------------------
MaSDH2       ADIVKIATTAVDIVDVTRMFQVIVHCQ---------------------------------
PvSDH2       ADIVKISTTALDITDSARLLQLLLHSQ---------------------------------
GmSDH1       ADVVKIATTALDITDCARLFQVLVHSQ---------------------------------
FvSDH2       ADIVKVATTALDITDNASIFQVLARSQ---------------------------------
```

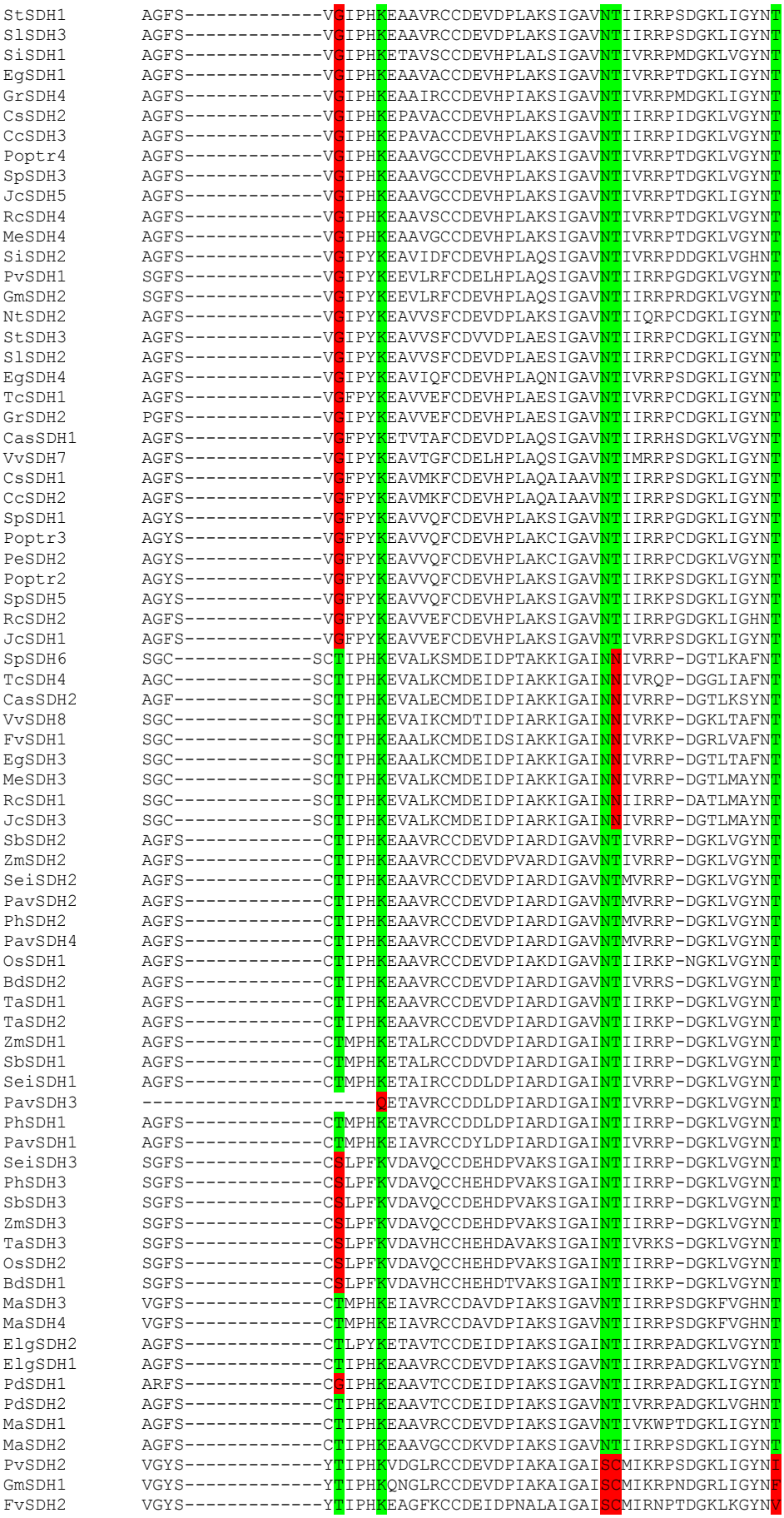
```
PpSDH1       ADIVKIATTALDITDSARVFQVLVHSQ---------------------------------
PmSDH3       ADIVKIATTALDITDSARVFQVLVHSQ---------------------------------
MdSDH1       ADIVKIATTGLDITDSARVFQVLAHSQ---------------------------------
PbSDH1       ADIVKIATTGLDITDSARVFQVLAHSQ---------------------------------
MdSDH4       ADIVKIATTALDITDSARVFQVLVHSQ---------------------------------
PbSDH4       ADIVKIATTALDITDSARVFQVLVHSQ---------------------------------
TcSDH3       ADIVKIATTALDIIDNARLFQVLVHSQ---------------------------------
GrSDH3       ADIVKIATTAVDIIDNARIFQVLVHSQ---------------------------------
CasSDH3      ADIVKIATTAQDITDSARIFQLLAHSQ---------------------------------
DkSDH        ADIVKIATTALDITDSARILQLIAHSQ---------------------------------
EgSDH2       ADIVKIATTALDISDCPRIFEVLAHSQ---------------------------------
VvSDH9       ADIVKIATTALDITDCARIFQVLAHSQ---------------------------------
SpSDH2       ADIVKIATTALDITDNARMFHIIANLQ---------------------------------
Poptr5       ADIVKVATTALDITDNARMFHIIVNLQ---------------------------------
PeSDH3       ADIVKVATTAIDITDNARMFHIIVNLQ---------------------------------
MeSDH1       ADIVKIATTALDITDNARMFQVLVHSQ---------------------------------
JcSDH4       ADIVKIATTALDIVDNARMFQVLVHSQ---------------------------------
CpSDH        ADIVKIATTALDITDVARIFQITVHSQ---------------------------------
AtSDH        ADIVKIATTAVDIADVARMFHITSKAQ---------------------------------
AlSDH        ADIVKIATTAVDIADVARMFHITSSAQ---------------------------------
PpSDH2       ADIVKIATTCLDITDVARTFQIAVHSQ---------------------------------
PmSDH2       ADIVKIATICLDITDVARTFQIAVHSQ---------------------------------
PmSDH1       ADIVKIATICLDITDVARTFQIAVHSQVS-------------------------------
MdSDH2       SDIVKIATTSLDITDVARTFQITVHSQ---------------------------------
PbSDH2       ADIVKIATASLDITDVARTFQITVHSQ---------------------------------
MdSDH5       ADIVKIATSAXDITDVARXFHITVHSQIIEEHCPFLMCGKXKGFSRGKGFVLFVGLTNKI
FvSDH3       ADIVRITTTALDITDVARIIQISVHSH---------------------------------
FsSDH        ADIVKFATTALEITDVARIFQIIVHSQVS-------------------------------
JrSDH        ADIVKIATTALEIADVARIFQITVHSQ---------------------------------
FvSDH4       ADIVKIATTALDITDVVRIFQITVHSQ---------------------------------
PpSDH3       ADIVKIATTAMDITDVARIFQITVHSQ---------------------------------
MdSDH3       ADIVKIATSALDITDVARIFHVTVHSQ---------------------------------
PbSDH3       ADIVKIATSALDITDVARIFHVTVHSQVS-------------------------------
NnSDH1       --------MPIYIIKVLFQYDTTLKFQ---------------------------------
NnSDH2       ADIVKIATTALDITDVVHMFQITVHSQ---------------------------------
SiSDH3       ADIVKFATTALDITDVVRVFQIMVHSQ---------------------------------
NtSDH1       ADIVKFATTALDIMDVARVFQITVHSQ---------------------------------
StSDH2       ADIVKFATTAQDITDVARVFQITVHSQ---------------------------------
SlSDH1       ADIVKFATTAQDITDVARVFQITVHSQ---------------------------------
MtSDH        ADIVKIATTAVEITDVARMFQIMVHSQ---------------------------------
PvSDH3       ADIVKIATTATDITDVARMFQIMVHSQ---------------------------------
GmSDH3       ADIVKIATTALDITDVARMFQIMVHSQVR-------------------------------
LuSDH1       ADIVKIATTALDITDNARMFRITVHSQ---------------------------------
LuSDH2       ADIVKIATTALDITDNARMFRITVHSQ---------------------------------
TcSDH2       ADIVKIVTTALEITDVARIFQITVHSQ---------------------------------
GrSDH1       ADIVKIATTAIDITDVARIFQITVHSQ---------------------------------
SpSDH4       ADIVKIATTALDISDVARIFQITVHSQVR-------------------------------
Poptr1       ADIVKIATTALDISDVARIFQITVHSQVR-------------------------------
PeSDH1       ADIVKIATTALDISDVARIFQITVHSQ---------------------------------
MeSDH2       ADIVKIATTASDIADVARVFQITVHSQ---------------------------------
RcSDH3       ADIVKIATTALDITDVARIFQITVHSQ---------------------------------
JcSDH2       ADIVKIVTTALDITDVARIFQITVHSQ---------------------------------
VvSDH5       ADIVKIATTALEITDVARIFQITVHSQ---------------------------------
EgSDH5       ADIVKIATTALDITDVARMFHITVHSQVS-------------------------------
CsSDH3       ADIVKFATTALDITDVARVFQITVHSQ---------------------------------
CcSDH1       ADIVKFATTALDITDVARVFQITVHSQ---------------------------------
```

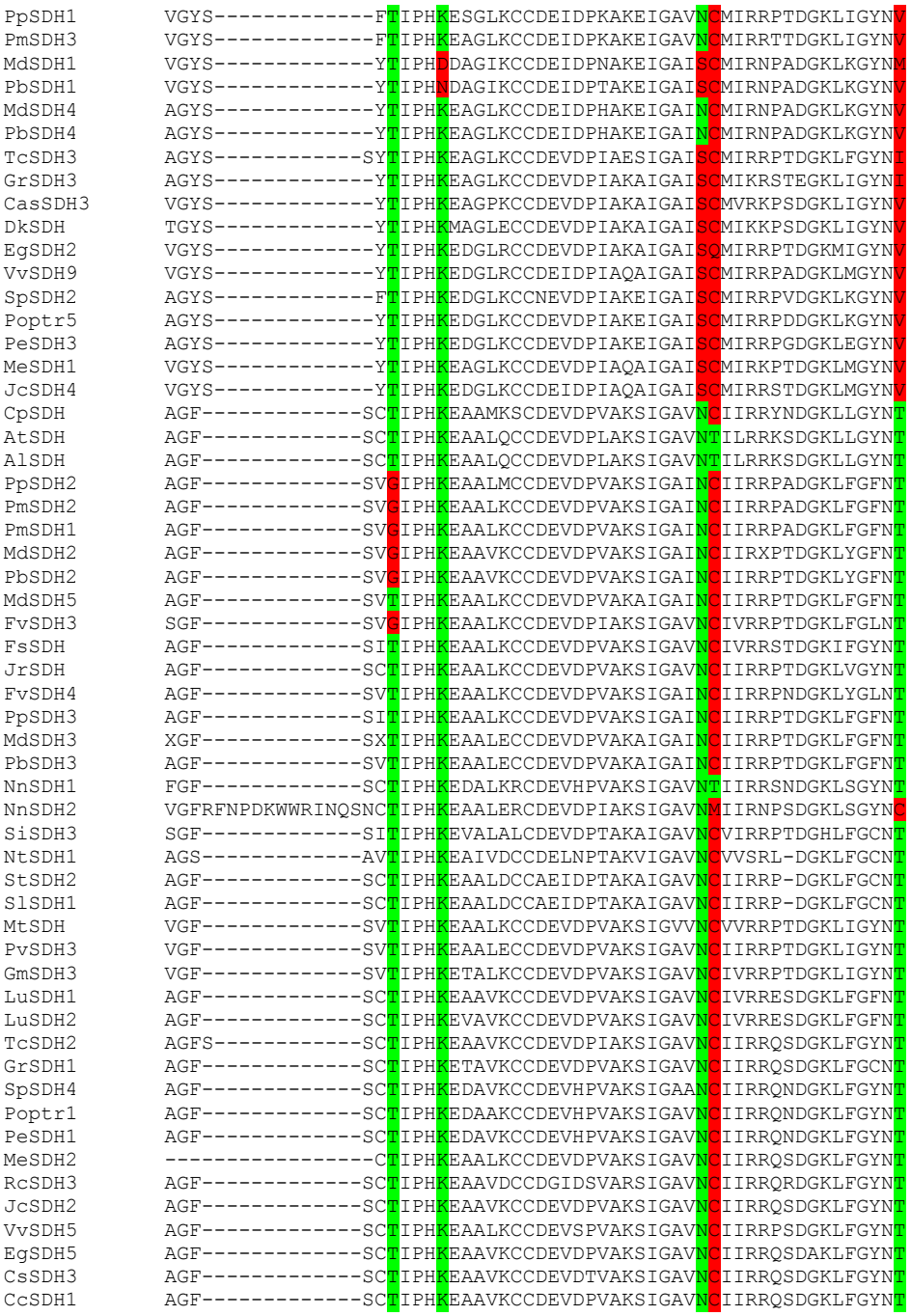
```
StSDH1       -------LPLIARAAGDRGLISQLLGPKYGAFFVCGSLAGK--STPGLPPLTSIKQVYKL
SlSDH3       -------VPLIVRAAGDRGLISQLLGPKYGAFFVCGSLGGK--YTPGLPSLTTIKQVYKL
SiSDH1       -------VPLIARAMGERGLISQLLGPKYGASMIFGSLGGK--FVPGLPPLVSIKHVYKL
EgSDH1       -------VPLIATTVGDRGLIGQLLGPKFGGFLAYGSLEG---AIPGLPTLTSLRQVYKI
GrSDH4       -------VPLIAMAVGNSGLISQLLGPKFGGFLVYGSLGGK--PVPGLPSLVSLRQVYKL
CsSDH2       -------VPLIALAVGSRGLISQLLGPKFGGFLVYGSLGGK--SVPGLPTLVSLKQVYQL
CcSDH3       -------VPLIALAVGSRGLISQLLGPKFGGFLVYGSLGGK--SVPGLPTLVSLKQVYQL
Poptr4       -------IPLIALAVGSRGLISQLLGPKFGGFLVYGSLSDK--AVPGMPTLLSLRQIYKL
SpSDH3       -------MPLIALAVGSRGLISQLLGPKFGGVLVYGSLSDK--AVPGMPTLLSLRQVYKL
JcSDH5       -------VPLIALAVGSRGLISQLLGPKFGGFLVYGSLEEK--PIPGMPTLFSLRHIYKL
RcSDH4       -------VPLIALAVGSRGLISQLLGPKFGGFLVYGSLRDK--AVPGMPTLFSLRHVYKL
MeSDH4       -------VPLIALAVGSRGLISQLLGPKFGGFLVYGSLGDK--TVPGMPTLFSLRHVYKL
SiSDH2       -------VPMIAYCTGERGLISQLLTFKFGGVFTYGSLQGT--SICGLPSINSLRQAYGV
PvSDH1       -------VPLIAYSVGERGLISHLLSPKFGGFFVYGSLAGN--PIPGMPSLESLKEAYKL
GmSDH2       -------VPLIAYSVGERGLISQLLSPKFGGFFVYGSLAGN--PIPGLPSLDSIQEAYKL
NtSDH2       -------VPLIAYSLGERGLISQLLGPKFGSVLVYGSLDCN--AVPGLPTLGSLRQAYGV
StSDH3       -------VPLIAYSVGERGLISQLLGPKFGSVLLYGSLDGN--AVPGLPSLASLRQAYGI
SlSDH2       -------VPLIAYSVGERGLISQLLGPKFGSVLLYGSLDGN--AVPGLPSLASLRQAYGV
EgSDH4       -------MPVVAYAVGERGLISQLLSPKFGSNLVYGSIEGM--AVPGLPTLESLRKAYKV
TcSDH1       -------VPIIAYSVGERGLISQLLCPKFGGFLAYGSIVGH--SLPGMPSLYSLRHTYKL
GrSDH2       -------MPVIAYSVGERGLISQLLCPKFGGFLAYGSIEGH--SIPGMPSLYSVKHTYKH
CasSDH1      -------IPLITYSIGDRGLISQLLGPKFDGFLIYGSIEGN--PVPGLPTLHNLRQAYGV
VvSDH7       -------MPLIAYSIGDRGFMSQILCRKFGGFLVYGSMEGS--PVAGLPTLESLREAYKV
CsSDH1       -------VPIIAYSVGERGLVSQLLSPKFNGALVYGSLKGT--PVLGLPTVESLRQTYKV
CcSDH2       -------VPIIAYSVGERGFVSQLLSPKFNGALVYGSLKGT--PVSGLPTVESLRQTYKV
SpSDH1       -------VPAVAYSVGERGLISQLLCPKFGGALVYGSMEGN--SIPGLPTLDSLKESYKV
Poptr3       -------VPAVAYSVGERGLISQLLCPKFGGALVYGSMEGN--SIPGLPTLDSLREAYKV
PeSDH2       -------VPAVAYSVGERGLISQLLCPKFGGALVYGSMEGN--SIPGLPTLDSLREAYKV
Poptr2       -------VPAVAYSLGERGLISQLLCPKFGGALVYGAMEGN--SIPGLPTLDSLREAYKV
SpSDH5       -------VPAVAYSLGERGLISQLLCPKFGGALVYGAMEGN--SIPGLPTLDNLRETYKV
RcSDH2       -------VPIIAYSVGERGLISQLLSPKFGGSLLYGSMEGS--SIPGLPTLDSLREAYKV
JcSDH1       -------VPLIAYSVGERGLISQILSPKFGGSLVYGSIEGN--LIPGLPTLDSLREAYKV
SpSDH6       -------VPIIGIVMGERGLISRLLSPKFGGYLTYGALEAGAISAPGQPTAKDLLDLYNF
TcSDH4       -------IPTIGVAMGEKGLVSQLLSPKFGGYLTYGALETGSI-APGQPTLKNLLDLYNF
CasSDH2      -------IPMIGIAMGERGLISRLLSPKFNGYLTYAALEAGAISAPGQPTVKDLLDLYNF
VvSDH8       -------VPTIAIVMGERGLISRLLSPKFGGYLTYGALEAGAISAPGQPTAKDLLDLYNF
FvSDH1       -------VPTIGIVMGERGLISRLLSPKFGGYLTYGALEAGAISAPGQPTARDLLDLYNF
EgSDH3       -------IPIIGIVMGERGLISRILSPKFGGYLTYGALEAGAISAPGQPTAKDLLDLYNF
MeSDH3       -------VPIIGIVMGERGLISRLLSPKFGGFLTYGALEAGGVSASGQPTAKDLLDLYNF
RcSDH1       -------VPVIGIVMGERGLISRLLSPKFGGYLTYGALEAGAISAPGQPTAKDLLDLYNF
JcSDH3       -------IPVIGIVMGERGQISRLLSPKFGGYLTYGALEAGAISAPGQPTVKDLLDLYNF
SbSDH2       -----KQVPIIGLVMNDRGFISRVLCPKYGGFLTFGSLKKGKESAPAQPTAADLINVYNI
ZmSDH2       -----KQVPIIGLVMNDRGFISRVLCPKYGGFLTFGSLKKGKESAPAQPTAADLINLYNI
SeiSDH2      -----KQVPIIGLVMNDRGFISRVLCPKYGGYLTFGSLEKGKESAAAQPTVADLINVYNI
PavSDH2      -----KQVPIIGLVMNERGFISRVLCPKYGAYLTFGSLEKGKESSPAQPTAADLINLYNI
PhSDH2       -----KQVPIIGLVMNERGFISRVLCPKYGAYLTFGSLEKGRESAPAQPTAADLINLYNI
PavSDH4      -----KQVPIIGLVMNERGFISRVLCPKYGAFLTFGSLEKGKESSPAQPTAADLINLYNI
OsSDH1       -----KQVPIIGLVMNDRGFISRVLCPKFGGYLTFGSLEKGKESAPAQPTVTDLINVYNI
BdSDH2       -----KQVPIIGLVMNDRGFISRVLCPKFGGYLTFGSLAKGKESAPAQPTAADLINVYNI
TaSDH1       -----KQVPIIGLVMNDRGFISRVLCPKFGGYLTFGSLEKGKESAPSQPTAADLINVYNI
TaSDH2       -----KQVPIIGLVMNDRGFISRVLCPKFGGYLTFGSLEKGKESAPSQPTAADLINVYNI
ZmSDH1       -----KQVPIIGLVRKEHGFISRIICAKYGGYLTFASLENGKESTTEQPTVADLINKYKI
SbSDH1       -----KQVPIIGLVMKEQGFISRILCAKYGGYLTFASLEKGKESAPGQPTVSDLINKYKI
SeiSDH1      -----KHVPIIGLVMKERGFISRVLCAKYGGYLTFASLEKGKESTPGQPTVADLITKYKI
PavSDH3      -----KQLPIIGLVMKERGFISQILCTKYGGYLTFASLEKGKESAPRQPTVAELVNKYKI
PhSDH1       -----KQLPIIGLVMKEQGFISRILCAKYGGYLTFASLEKGKESAPGQPTVAELINKYEI
PavSDH1      -----KQLPIIGLVMKERGFISQILCAKYGGYLTFASLEKGKESAPGQPTVTELINKYKI
SeiSDH3      -------VPMIGLVMSEKGLMSRVLAPKFGGYLTFGILNAAKTSASGQPTVEELLEIYNI
PhSDH3       -------VPMIGLVMSERGLMSRVLAPKFGGYLTFGILNAAKTSASGQPTVEELLDVYNI
SbSDH3       -------VPMIGLVMSERGLMSRVLAPKFGGYLTFGILNATKTSASGQPTVEDLLNIYNI
ZmSDH3       -------VPMIGLVMSEKGLMSRVLAPKFGGYLTFGILNATKTSASGQPTVEDLLDIYNI
TaSDH3       -------VPMIGLVMSERGLMSRVLSPKFGGYLTFGILDATKTSASGQPTLEELLDVYNI
OsSDH2       -------VPMIGLVMGEKGLMSRVLSPKFGGYLTFGTLDATKISAPGQPTVKELLDIYNI
BdSDH1       -------VPMIGLVMSERGLMSRVLSPKFGGYLTFGILDTTKTSASGQPTVEELLDIYNI
MaSDH3       -------VPMIGLVMGERGLLSRVLCPKYGGYLTFGTLGEGKASAPGQPTVSQLLHLYNI
MaSDH4       -------VPMIGLVMGERGLLSRVLCPKYGGYLTFGTLGEGKASAPGQPTVSQLLHLYNI
ElgSDH2      -----KQVPMIGFVMGERGVISQLLCPKFGGYLTFGALDAGKESAPRQPTLSDLLNVYNI
ElgSDH1      -------VPVIGLVMGERGMIFRLLCPKFGGYLTFGSLDVGKESAPGQPTISDLLNVYNI
PdSDH1       -------VSMIGFVMGERGVVSRLLCPKFGGYLTFGTLDAGKESAPGQPTISDLRNVYNI
PdSDH2       -----KQVPMIGLVMGERGVISRLLCPKFGGYLTFGTLDAGKESAPGQPTISDLRNVYNI
MaSDH1       -------VPMIGLVMEEKGLISRLLCPKYGGYLTFATLGAGKESAPGQPTISDLLNVYNF
MaSDH2       -------VPMIGLVMGERGLISRLLCPKFGGYLTFGTLGVGKESAPGQPTISDLLNVYNI
PvSDH2       -------VPMIGIAIGEKGILSRILCAKFGGFLTFGSIEDGAVSVPGQPTIKELLDLYNF
GmSDH1       -------VPMIGIAMGEKGFISRVLCAKFGGFLTFGSIEAGAISAPGQTTIKELLDLYNF
FvSDH2       -------VPMIGLVMGDKGLISRVLSAKYGSFATFGTLVAGEVSAPGQPTVTDLLELYNF
```

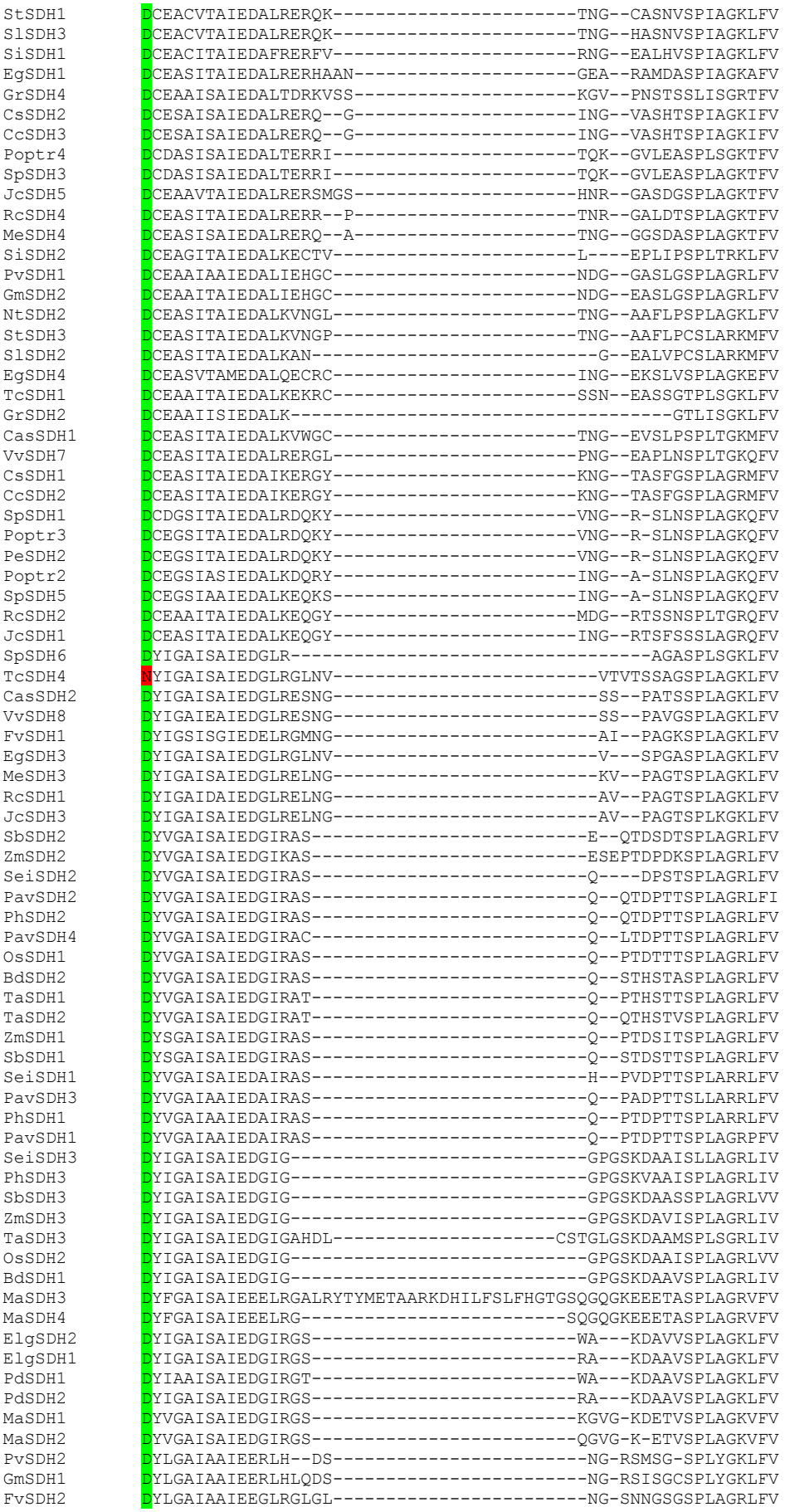
```
PpSDH1       -------VPMIGLVMGEKGLISRVLSAKYGAFLTFGTIEAGVVSAPGQPTVKDLLDLYNF
PmSDH3       -------VPMIGLVMGEKGLISRVLSAKYGAFLTFGTIEAGVVSAPGQPTVKDLLDLYNF
MdSDH1       -------VPMIGLVMGEKGLISRVLSAKYGAFLTFGTIEAGVVSAPGQPTVRDLLDLYNF
PbSDH1       -------VPMIGLVMGEKGLISRVLSAKYGAFLTFGTIEAGVVSAPGQPTVRDLLDLYNF
MdSDH4       -------VPMIALVMGEKGLISRVLSAKYGAFLTFGTIEAGVVSAPGQPTVRDLLDLYNF
PbSDH4       -------VPMIALVMGEKGLISRVLSAKYGAFLTFGTIEAGVVSAPGQPTVRDLLDLYNF
TcSDH3       -------VPMIGLVMGERGLMSRILAAKFGGFLTFGTLEEGVISAPGQPTVKELLDLYNM
GrSDH3       -------VPMIGLVMGERGLMSRILAAKFGGFLTFGSLEAGLVSAPGQPTVKELLDLYNM
CasSDH3      -------VPTIGIVMGERGLMSRILCAKFGGFLTFGALESGIVSAPGQSTLKDLLELYNF
DkSDH        -------VPTIGLAMGERGLISRILCAKFGGFLTFGALESGIHSAPGQPTLRDLLDLYNF
EgSDH2       -------VPTIAIAMGERGLISRILAAKFGGYLTFGAIEAGVVSAPGQPSIKDLLDLYNL
VvSDH9       -------VPTIGIAMAERGLISRILSAKFGSYLTFGSLEAGVVSAPGQPTVKDLLDLYNF
SpSDH2       -------VPVIGLVMGERGLMSRVLAAKYGAFLTFGSIEAGVVSAPGQPTIKDLLELYNF
Poptr5       -------VPMIGLVMGERGLMSRVLAAKYGAFLTFGSIEAGVVSAPGQPTVKDLLELYNL
PeSDH3       -------VPMIGLVMGERGLMSRVLAAKYGGFLTFGSIEAGVVSAPGQPTVKDLLELYNL
MeSDH1       -------VPMIGLVMGERGLISRILAAKYRGFLTFGSIEAGVVSAPGQPSIKELLDLYNI
JcSDH4       -------VPMIGLVMGERGLMSRILAAKFGGFLTFGSIEAGVASAPGQPTIKDLLDLYNF
CpSDH        -------VPVIGLVMGERGLISRILCAKFGGYLTFGTLEAGIVSAPGQPTIKDLLYLYNF
AtSDH        -------VPTIGLVMGERGLMSRILCSKFGGYLTFGTLDSSKVSAPGQPTIKDLLDLYNF
AlSDH        -------GPTIGLVMGERGLMSRILCSKFGGYLTFGTLDSSKVSAPGQPTIKDLLDLYNF
PpSDH2       -------VPLIGLGLGERGLISRILSAKFGGYLTYGMLEPGTVSAPSQPTIKDLIHLYNF
PmSDH2       -------VPLIGLGLGERGLISRILSAKFGGYLTYGVLEPGTVSAPSQPTIKDLIHLYNF
PmSDH1       ------RVPLIGLGLGERGLISRILSAKFGGYLTYGVLEPGTVSAPSQPTIKDLIHLYNF
MdSDH2       -------VPFIGLVLGERGLISRILSAKFGGYLTYGRLESGIISAPTQPTVKDLLHLYNF
PbSDH2       -------APFIGLVLGERGLISRILSAKFGGYLTYGTLESRIVSAPTQPTVKDLLHLYNF
MdSDH5       DSFMEFQVPTIGLVMGERGLISRILCAKFGGYLTFGTLDSXIVSAPGQPTMKDILHLYNF
FvSDH3       -------VPVIGLAMGERGLISRILCAKFGGYFTSGTLHSGIVSVPGEPAIKDILDVYNF
FsSDH        ------SVPIIAIVMGERGFISRILCPKFGGFLSFGTIESGIVSAPGQPIMKDLLHLYNF
JrSDH        -------VPIIGIVMGERGFMSRILCPKFGGFLTFGTIESGIVSAPGQPTMKDLLHLYNL
FvSDH4       -------VPIIGLVMGERGLISRILCAKFGGYLTFGTIDSGAVSAPGQPTIKDILNLYNF
PpSDH3       -------VPTIGLVMGERGLISRILCAKFGGYLTFGTLDSGIVSAPGQPTMKDILHLYNF
MdSDH3       -------VPTIGLVMGERGLISRILCAKFSGHLTFGTLDSGIVSAPGQPTIKDILHLYNF
PbSDH3       ------SVPTIGLVMGEKGLISQILCAKFGGFLTFGSLDSGIVSAPGQPTIKNILHLYNF
NnSDH1       -------VPIIGLVMGERGLMSRILCPKFGGYLTFGTLEAGKISAPGQPTIKDLLDLYNF
NnSDH2       -------VPIIGLVMGERGLISRVLCPKFGGYLTFGTLEAGKLSAPAQPTIKELLNFYNF
SiSDH3       -------IPIIAMVMGERGLISRILCPKFGGYLTFGTLEPGKVSAPGQPTIDELVNIFNL
NtSDH1       -------VPIIAMVMGEKGLMSRILCPKFGGYLTFGTLEVGKVSAPGQPTIKDLLNIYNF
StSDH2       -------VPIIAMVMGEKGLMSRILCPKFGGYLTFGTLEVGKVSAPGQPTVEDLLNLYNF
SlSDH1       -------VPIIAMVMGEKGLMSRILCPKFGGYLTFGTLEVGKVSAPGQPTVEDLLNLYNF
MtSDH        -------VPFIGLVMGDRGLISRILCAKFGGYLTFGTLESGVVSAPGQPTLKDLLHLYNF
PvSDH3       -------VPFIGLVMGDRGLISRILCAKFGGYLTFGTLESGVVSAPGQPTLQDLLYLYNL
GmSDH3       ------RVPFIGLVMGDRGLISRILSAKFGGYLTFGTLESGVVSAPGQPTLKDLLYLYNL
LuSDH1       -------VPIIGLVMGERGLMSRVLCGKFGGYLTFGTLESGVVSAPGQPLIKDLLDLYNF
LuSDH2       -------VPIIGLVMGERGLMSRVLCGKFGGYLTFGTLESGVVSAPGQPLIKDLLDLYNF
TcSDH2       -------VPIIGLVMGERGLISRILCAKFGGYLTFGTLEAGVVSAPGQPTINDLLNLYNF
GrSDH1       -------VPIIGLVMGERGLISRILCTKFGGYLTFGTLEGGVVSAPGQPTINDLLNLYNF
SpSDH4       ------SVPIIGLVMGERGLISRILCAKFGGYLTFGTLESGVVSAPGQPTIKDLLDLYNF
Poptr1       ------SVPIIGLVMGERGLISRILCAKFGGYLTFGTLESGVVSAPGQPTIKDLLDLYNF
PeSDH1       -------VPIIGLVMGERGLISRILCAKFGGYLTFGTLESGVVSAPGQPTIKDLLDLYNF
MeSDH2       -------VPIIAMVMGERGLISRVLCAKFGGYLTFGTLESGIVSAPGQPTIKDLLDLYNF
RcSDH3       -------VPIIGIVMGERGLISRILCAKFGGYLTFGTLESGIVSAPGQPMIKDLLDLYNF
JcSDH2       -------VPIIGLVMGERGLISRILCAKFGGYLTFGTLESGIVSAPGQPMIKDLLDLYNF
VvSDH5       -------VPVIGLVMGERGLISRILCPKFSGYLTFGSLEPGIVSAPGQPTIKDLLNLYNF
EgSDH5       ------SVPVIAMVMGERGLISRVLCAKFGGFLTFGTLESGVVSAPGQPTIKDLLDLYNF
CsSDH3       -------VPIIGLVMGERGLISRILCAKFGGFLTFGTLENGIVSAPGQPTIKDLLDLYNF
CcSDH1       -------VPIIGLVMGERGLISRILCAKFGGFLTFGTLENGIVSAPGQPTIKDLLDLYNF
```

```
StSDH1       QYVNPDTRVFGVISNPVGHSKGPLLHNPAFRHTGYNGIYVPLLVDNVKEFFRVFSCNDY-
SlSDH3       QYVNPDTRIFGVISNPVGHSKGPLLHNPAFRHTGYNGIYVPLLVDNVKEFFRVFSCNDY-
SiSDH1       ENMNADTKIFGVVSNPVGHSKGPLLHNPAFRQTGYNGIYVPLLVDDIKEFFRVYSCNDF-
EgSDH1       EHMNADTKVFGLISNPVGHSKGPLLHNPTFRYVGFNGIYVPMLVDNIKEFFKVYTSQDF-
GrSDH4       EYTNVDTKVFGLVSNPVAHSKGPILYNPTFRHMGYNGIYVPMLVDDIEEFFETYSGSDF-
CsSDH2       EHINPDTKIFGLVSNPVGHSKGPILHNPAFRHTRFNGIYVPMLVDDVKEFFRTYSGTDF-
CcSDH3       EHINPDTKIFGLVSNPVGHSKGPILHNPAFRHTRFNGIYVPMLVDDVKEFFRTYSGTDF-
Poptr4       EYINADTKVFGLISNPVGHSKGPVLHNPAFRHTGYNGIYVPMQVDDVKEFFRTYTSSDF-
SpSDH3       EYINADTKVFGLISNPVGHSKGPILHNPAFRHTGYNGIYVPMQVDDVKEFFRTYTSSDF-
JcSDH5       EYINPETKVFGLISNPVGHSKGPILHNPAFRHTGYNGIYVPMQVDDIEEFFRTFTSTDF-
RcSDH4       EYINSDTKVFGLISNPVGHSKGPILHNPAFRHTGYNGIYVPMQVDDIKEFFRTYTSTDF-
MeSDH4       EYINAETKVFGLISNPVGHSKGPILHNPAFRHTGYNGIYVPMQVDDIKAFFRTYTGTDF-
SiSDH2       KNIDKNTVVFGLISKPVSHSKGPILHNPAFRHVYYNGIYVPMLVDDLKAFFTVYTSADF-
PvSDH1       ERVNEDTKVFGLISKPVSHSKGPILHNPSFRHVNYNGIYVPMFVDDLKEFFSTYPSPDF-
GmSDH2       EHVNADTKVFGLISKPVSHSRGPILHNPSFKDVNYNGIYVPMFVDDLKKFFSTYPSPDF-
NtSDH2       DFMDTDTKVFGLISKPVGHSKGPILHNPTFRHVGYNGIYVPMFVDDLKEFFRVYSSPDF-
StSDH3       NFMDNDTKVFGLISKPVGHSKGPILHNPTFRHVGYNGIYVPMFVDDLKEFFRVYSSPDF-
SlSDH2       DLMDNDTKVFGLISKPVGHSKGPILHNPTFRHVGYNGIYVPMFVDDLKEFFRVYSSPDF-
EgSDH4       EHINSDTKVFGLISKPVGHSKGPILHNPTLRHMNFNGLYVPMFVDNLKEFFEVYSTPDF-
TcSDH1       DYLNSETKVFGLVSKPVGHSKGPLLHNPTLRHENFNGVYVPMFVDNLKEFFSIYSSPDF-
GrSDH2       DLMNSETKVFGLVSKPVSHSKGPILHNPTLRHENFNGVYVPLFVDNLKEFFSTYSGSDF-
CasSDH1      DYLNAETKVFGLISKTSGPCQRSPSPYPTFRHVGYNGIYVPMLVDDLKEFFSVYSSPDY-
VvSDH7       QYINKDTKVFGLISKPVGHSKGPILHNPAFRHVNYNGIYVPMLVDDLKEFFSIYSSPDF-
CsSDH1       EHINADTKVFGLISKPVGHSKGPILHNPTFRHVNYNGIYVPMFVDDLKKFFSTYSSPDF-
CcSDH2       GHINADTKVFGLISKPVGHSKGPILHNPTFRHVNYNGIYVPMFVDDLKKFFSTYSSPDF-
SpSDH1       DCIDSDTKVFGLVSKPVSHSKGPLLHNPTLRHAIFNGIYVPMFVDDLKKFFDVFASPDF-
Poptr3       DCINSDTKVFGLVSKPVGHSKGPLLHNPTLRHVNFNGIYVPMFVDDLKKFFDVYASPDF-
PeSDH2       DCINSDTKVFGLLSKPVGHSKGPLLHNPTLRHVNFNGIYVPMFVDDLKKFFDVYASPDF-
Poptr2       ENINSDTKVFGLVSKPVSHSKGPILHNPAFRHANFNGIYVPMFVDDLKEFFEVYASPDF-
SpSDH5       ENINSDTKVFGLVSKPVSHSKGPILHNPAFRHANYNGIYVPMFVDDLKEFFEVYTSPEF-
RcSDH2       AYINSDTKVFGLVSKPVSHSKGPLLHNPTFRHANYNGTYVPMFVDDLKEFFSVYSSPDF-
JcSDH1       ERINSDTKVFGLISKPVGHSKGPILHNPTFRHVNYNGIYVPMFVDDLKEFFSVYSSPDY-
SpSDH6       RLIRPNTKVYGINGKPVSHSKNPLLFNAAFKAVGLNAVYLHFLVDDVEEFFNTYSCIDFA
TcSDH4       RLITPDTKVYGIIGKPVGHSKSPLIFNAAFKSVGLNAVYVHFLVDDVEMFFNTYSSPDF-
CasSDH2      RLIRPDTKVYGIIGKPVGHSKSPLLFNESFKKVGLNSVYVHLLVDDVKKFFNTYSSVDF-
VvSDH8       RLVKPDTKVYGIIGKPVGHSKSPLLFNAAFKKVGLNAVYVHLLVDDVEKFFNTYSAPDFI
FvSDH1       RLIRPDTKIYGIIGKPVGHSKSPLLFNAAFKSVGLNAVYVHFLVDDVEKFFNTYSSVDF-
EgSDH3       RLIRPDTKVYGIIGKPVGHSKSPLLFNASYKSVGLNAVYLHFLVDDVEKFFNTYSAVDFA
MeSDH3       RLIRPDTKVYGIIGKPVGHSKSPLLFNASFKSVGLNAVYMHFLVDDVEKFFNTYSSVDFA
RcSDH1       RLIRPDTKVYGIIGKPVGHSKSPLLFNAAFKSVGLNAVYVHFLVDDVEKFFSTYSSVDFA
JcSDH3       RLIRPDTKVYGIIGKPVGHSKSPLLFNAAFKSVGLNAVYMHFLVDDVEKFFNTYSSVDFA
SbSDH2       RQIGPDTKVFGIIGKPVGHSKSPILHNEAFRSVGSNAVYVPFLVDDLAKFLNTYSSPDF-
ZmSDH2       RQIGPDTKVFGIIGKPVGHSKSPILHNEAFRSVGFNAVYVPFLVDDLAKFLDTYSSPDF-
SeiSDH2      RQIGPDTKVFGIIGKPVGHSKSPILHNEAFRSVGFNAVYVPFLVDDLAKFLNTYSSPDF-
PavSDH2      RQIGPDTKVFGIIGKPVGHSKSPILHNEAFRSVGFNAVYVPFLVDDLAKFLNTYSSPDF-
PhSDH2       RQIGPDTKVFGIIGKPVGHSKSPILHNEAFRSVGFNAVYVPFLVDDLAKFLNTYSSPDF-
PavSDH4      RQIGPDTKVFGIIGKPVGHSKSPILHNEAFRSVGFNAVYVPFLVDDLAKFLNTYSTPDF-
OsSDH1       KQIGPDTKVYGIIGKPVGHSKSPILHNEAFRSVGLNAVYVPFLVDDLANFLNTYSSPEF-
BdSDH2       RQIGPDTRVFGIIGKPVGHSKSPILHNEAFRSVGMDAVYVPFLVDDLARFLTTYSSPDF-
TaSDH1       RQIGPDTKVFGIIGNPVGHSKSPILHNEAFRSVGLNAVYVPFLVDDLAKFLSTYSSPDF-
TaSDH2       RQIGPDTKVFGIIGNPVGHSKSPILHNEAFRSVGLNAVYVPFLVDDLAKFLSTYSSPDF-
ZmSDH1       RQIGPDTKVFGIIGNPVSHSKSPIVQNQAFRSVGFDAVFLPFLSDDLVQFLYTFSAPDY-
SbSDH1       RQIGPDTKVFGIIGNPVNHSKSPIVQNQAFRSVGFDAVFLPFLSDDLVEFLSTFSAPDY-
SeiSDH1      RQIGPDTKVIGIIGNPVSHSKSLIVQNQAFRSVGFNAVFLPFLSDDLVKFLKTFSSPEY-
PavSDH3      RQIGPDTKVFGIIGNPVGHSNSPIVQNQAFRSVGFDAVFLPFLSDD--------------
PhSDH1       RQIGPDTKVFGIIGNPVGHSKSPFLQNQAFRSVGFDAVFLPFLSDDLVKFLNTFSSPDY-
PavSDH1      KQIGPDTKVFGIIGNPVGHSKSPILQNQAFRSVGFDAVFLPFLSDDLVKFLSTFSSPDY-
SeiSDH3      RRIGPDTKVLGLIANPVKQSKSPILHNKCLQSVGYNAVYLPLLGDNLASFLETYSSPDF-
PhSDH3       RRIGPDTMVLGLIANPVKQSKSPILHNKCLQYTGYNAVYLPLLGDNLASFLETYSSPDF-
SbSDH3       KCIGPDTKVLGLIANPVKQSKSPILHNKCLQSIGYNAVYLPLLGDNLASFLETYSSPDF-
ZmSDH3       KCIGPDTKILGLIANPVKQSKSPILHNKCLQSIGYNAVYLPLLGDNLASFLETYSSPDF-
TaSDH3       RCIGPDTKVLGLIANPVKQSKSPILHNKCLQSVGYNAVYLPLLADDLARFLDTYSSPDF-
OsSDH2       RRIGPDTKVLGLIANPVKQSKSPILHNKCLQSIGYNAVYLPLLADDLARFLSTYSSPDF-
BdSDH1       RSIGPDTKVLGLVANPVKQSKSPILHNKCLQSVGYDAVYLPLLADDLARFLDTYSSPDF-
MaSDH3       RQIGADTKLTGLIGNPVRQTKSHIFHNAAFKSAGFDAVYVPLLVDDLTAFLHAFSSPDS-
MaSDH4       RQIGADTKLTGLIGNPVRQTKSHIFHNAAFKSAGFDAVYVPLLVDDLTAFLHAFSSPDS-
ElgSDH2      RKIRADTKVYGLIGKPVGQSKKPILHNAAFKSVGFNAVYVPLLVDDLANFLSAYCSPDF-
ElgSDH1      RKITADTKVYGIIGKPVGHSKGPILHNAAFKSVGLNAVYVPLLVDDLANFLSAYSSPDF-
PdSDH1       RKIRADTKVYGIVGKPVGHSTSPILHNAAFNSVGFNAVYVPLLVDDLANFLGAYCSPDF-
PdSDH2       RKIRADTKVYGIVGKPVGHSKSPILHNAAFKSVGFNAVYVPLLVDDLANFLSAYCSPDF-
MaSDH1       RQIGTDTKVFGIIGKPVGHSKSPILHNAAFKSVGLNAVYVPFLVDVLADFLNAYSSADF-
MaSDH2       RQIGADTKIFGIVGKPVGHSKSPVLLNAAFKSVGLDAVYVPLLVDDLVNFLNAYSSTDF-
PvSDH2       RHIGVDTKVHGVIGNPTSHSKSPHLYNAAFESVGFDGVYLPFLVDNVSDFLNTYSSLDF-
GmSDH1       RQIGVGTKVHGVIGNPIGHSKSPHLYNPAFKSVGFDGIYLPLLIDNVSDFLNTYSSPDF-
FvSDH2       RQIGADTKVHGVIGNPIGHSKSPHLYNAAFKANNFNGIYLPLLVDSVANFINTYNSPDW-
```

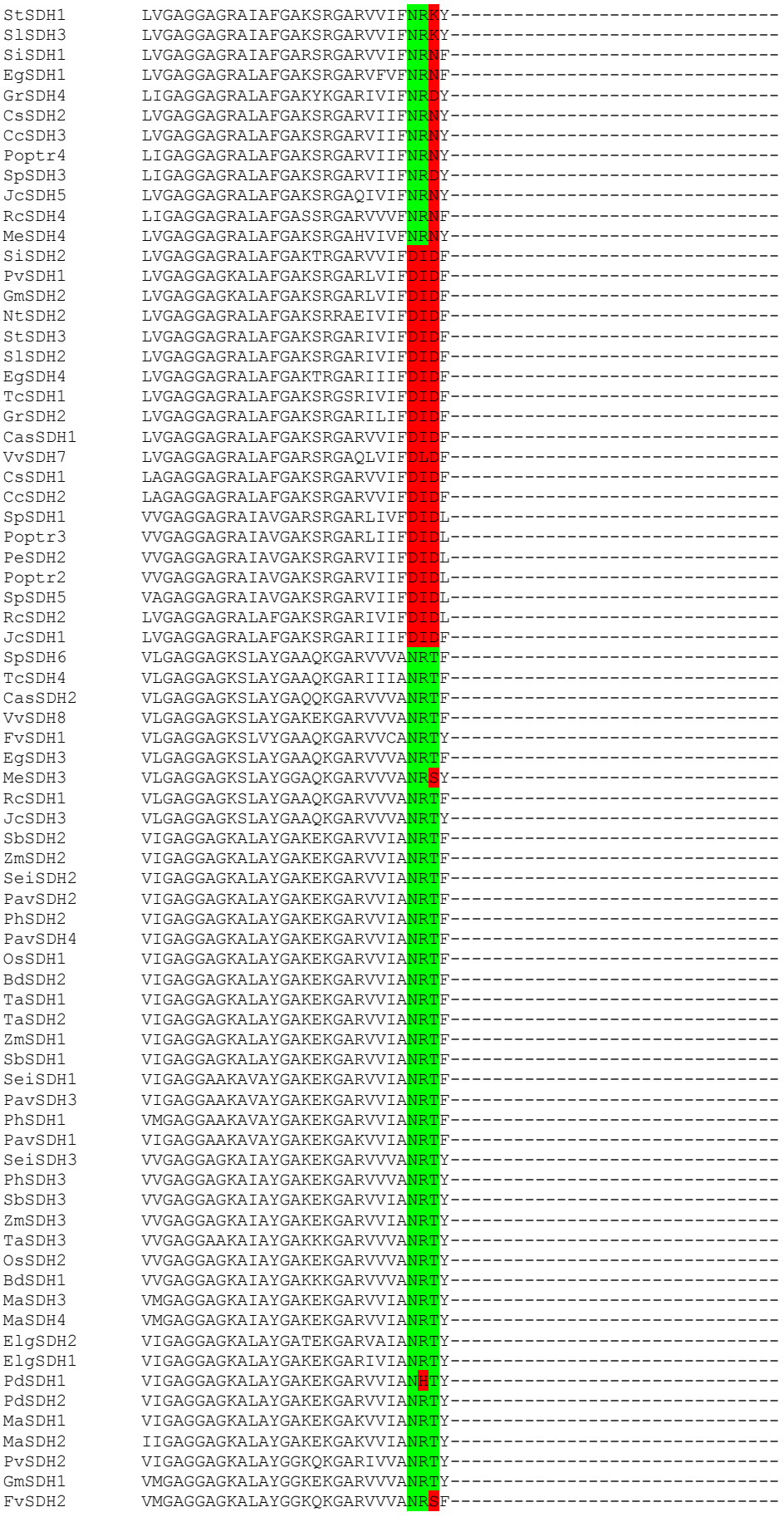
```
PpSDH1        RLIGLDTKVHGVIGNPIGHSKSPHLYNAAFKSINFNGIYLPLLVDSVANFIDTYNSPDF-
PmSDH3        RLIGPDTKVHGVIGNPIGHSKSPHLYNAAFKSINFNGIYLPLLVDSVANFIDTYNSPDF-
MdSDH1        RLIGADTKVHGVIGNPIGHSKTPHLYNAAFRSINFNGIYLPLLVDNVANFIDTYNSPDF-
PbSDH1        RLIRADTKVLGVMGNPISHSKTPHLYNAAFRSINFNGIYLPLLVDNVANFIDTYNSPDF-
MdSDH4        RIIGADTKVHGEIGNPIGHVKXPLLYNAAFRSINFNGIYLPLLVDNVANFIDTYNSPDF-
PbSDH4        RLIGADNKVYGVIGNPIRHSNGPHLYNAAFKSINFNGIYLPLLVDNVANFIDTYNSPDF-
TcSDH3        RHIGPDTKVHGVIGNPIDHSKSPHLYNAAFKSTGFNGIYLPLLVDNVSNFITTYSSPDF-
GrSDH3        RLIGPDTKVYGVIGNPIGHSKSPHIYNASFKSSSFNGIYLPLLVDDVSTFITTYSAPDF-
CasSDH3       RQIGPDTKVHGVIGNPIGHSKSPHLYNTAFRSAGFNGIYLPLLVDSVANFFNTYSSPDF-
DkSDH         RQIEPDTKVHGVIGNPIGHSKSPHLYNTAFKKAAFNGIYLPLLVDNVANFLNTYSSPDF-
EgSDH2        RQLGPDTKVHGVIGKPIGHSKSPHLYNAAFRSVNFNGIYLPLLVDNVANFINAYSSPDF-
VvSDH9        RQIGPDTKVHGVIGKPIGHSKSPHLYNSAFKSVGFNGIYLPLLVDSVKNFLATYSSPDF-
SpSDH2        RQIGADTKVYGVIGNPVGHSKSPLLYNAAFKHVGFNGIYLPLLVDSDANFISTYSSPDF-
Poptr5        RQIEADTKVHGVIGNPIGHSKSPHLYNAAFKSVGFNGIYLPLLVDSVANYISTYSSPDF-
PeSDH3        REIEADIKVHGVIGNPIGHSKSPHLYNAAFKSVGFNGIYLPLLVDSVANYISAYSSPDF-
MeSDH1        RLIGADTKIHGVIGNPIGHSKSPHLYNAAFKSVCFNGIYLPLLVDNVANYISTFSSPDF-
JcSDH4        RQIGADTKVHGVIGKPIGHSKSPHLYNAAFKSVGFNGIYLPLLVDNVANYISTYSSPDF-
CpSDH         RQIGPDTKVYGIIGKPVSHSKSPALHNEAFKSVGFNGVYVHLLVDDIANFLWTYSSPDF-
AtSDH         RRIGPDTKVYGIIGKPVSHSKSPIVHNQAFKSVDFNGVYVHLLVDNLVSFLQAYSSSDF-
AlSDH         RRIGPDTKVYGIIGKPVSHSKSPIVHNQAFKSVDFKGVYVHLLVDNLASFLQAYSSSDF-
PpSDH2        RQIGPYTKVFGVISKPVSHSKSPKLYNEVFKSVGFDGVYAHLLVDDIQKFFQTYSSMDF-
PmSDH2        RQIGPHTKVFGVISKPVSHSKSPKLYNEVFKSVGFDGVYAHLLVDDIQKFFQTYSSLDF-
PmSDH1        RQIGPHTKVFGVISKPVSHSKSPKLYNEVFKSVGFDGVYAHLLVDDIQKFFQTYSSLDF-
MdSDH2        RQIGPDTKVFGVISKPVSHSKSPKLYNEVFNSVGFNGVYAHLLVDDIAKFFQTYSCMDF-
PbSDH2        RQIGPDTKVFGVISKPVSHSKSPKLYNEVFNSVGFNGVYAHLLVDDIAKFFQTYSCMDF-
MdSDH5        RQIGTDTKVFGIIGKPVSHSKSPILYNEAFKSVGFNGVYVHLLIDDIVNFLHTYSSVDF-
FvSDH3        RLIRPCTKMFGVIGKPISHSKSPLLYNEGFKSVGFDGVYLHLLVDDIANFLHTYSSTDF-
FsSDH         RQIGPDTKVFGIIGKPVGHSKSPILYNEAFKSVGFNGVYVLSWWMT-CKISPDLLILRF-
JrSDH         RRIGPDTKVFGIIGKPVHHSKSPILYNEAFKSVCFNGVYIPLLVDDIANFLQTYSSTDF-
FvSDH4        RQIGPDTKVFGIIGKPVSHSKSPILHNEGFKSVGFNGVYVHLLVDDIANFLKTYSSADF-
PpSDH3        AQLGPDTKVFGIIGKPVSHSKSPFLYNEGFKSVGFNGVYVHLLVDDIAKFLQTYSSMDF-
MdSDH3        RQIGPDTKVFGIVGKPVSHSKSPMLYNEAFKSVCFNGVYVHLLVDDIAKFLHTYSSKDF-
PbSDH3        RQIGPDTKVFGIIGKPVSHSKSPMLYNEAFKSVCFNGVYVHLLVDDIAKFLHTYSSKDF-
NnSDH1        RQLGPDTKIYGLLGKPVSHSKGPFLFNAAFKAASFNGVYMPLLVDDVAKFFDVYSSPDF-
NnSDH2        RQLGPDTKVFGLIGNPVGHSKGPLLYNTAFKSIGFNGVYVPFLVDDVAKFFDAYSSPDF-
SiSDH3        RLIDTDTRVFGIIGKPVSHSKSPKLYNAAFKAAGFNGVYVHLLVDDVPKFFETYSSTDF-
NtSDH1        RQLGPDTRIFGIIGKPVSHSKSPLLYNEAFRSVGFNGVYMPLLVDDVANFFRTYSSLDF-
StSDH2        RQLGPDTKIFGIIGKPVSHSKSPLLYNESFRSVGFNGVFMHLLVDDIANFFRTYSSLDF-
SlSDH1        RQLGPDTKIFGIIGKPVSHSKSPLLYNESFRSVGFNGVFMHLLVDDIANFFRTYSSLDF-
MtSDH         RQIGPNTKVFGIVGKP---------------TIGFDGVFVFLLVDDLANFLRTYSSTDF-
PvSDH3        RQVAPDTKVYGIIGKPVGHSKSPILFNEVFKSVGYNGVYVFLLVDDLANFLRTYSSTDF-
GmSDH3        RQLAPDTKVFGIIGKPVGHSKSPILFNEVFKSIGLNGVYLFLLVDDLANFLRTYSSTDF-
LuSDH1        RQIQADTKVFGIIGKPVGHSKSPVLYNEAYRSVGFNGVYVHLLVDDLASFLKTYSSPDF-
LuSDH2        RQIQADTKVFGIIGKPVGHSKSPVLFNEAYRSLGFNGVYVHLLVDDLASFLKTYSSPDF-
TcSDH2        RQLGPDTKVYGVIGKPVGHSKSPMLYNEAFKSAGFNGVYVHLLVDDLSTFLQSYSSTDF-
GrSDH1        RQLGPDTKVYGIIGKPVGHSKSPMLYNEAFKSAGFNGVFVHLLVDDLEMFLRTYSSTDF-
SpSDH4        RQIGPDTKVFGIIGKPVGHSKSPVLYNEAFKSVGFNGVYVHLLVDDIARFLQTYSSTDF-
Poptr1        RLIGPDTKVFGIIGKPVGHSKSPVLFNEAFKSVGINGVYVHLLVDDIARFLQTYSSTDF-
PeSDH1        RQIGPDTKVFGIIGKPVGHSKSPLLFNEAFKSVGINGVYVHLLVDDIARFLQTYSSTDF-
MeSDH2        RLIGPDTKVYGIIGKPVGHSKSPTLYNEAFKSVGFNGVYMHF------------------
RcSDH3        RQIGPDTKVYGIIGKPVGHSKSPVLYNEAFKSVGFNGVYVHLLVDDIASFLQTFSSTDF-
JcSDH2        RQIGPDTKVYGIIGKPVGHSKSPVLYNEAFESVGFNGVYVHLLVDDLANFLRTFSSTDF-
VvSDH5        RQLGPDTKVFGVIGKPVSHSKSPHLYNEAFKSVGFNGVYVHLLVDDIAKFFHTYSAADF-
EgSDH5        RSIGPDTKVFGIIGKPVGHSKSPLLYNQAFKSAGFDGVFLHLLVDDVASFLQTYSSTDF-
CsSDH3        RQMGPDTKVFGIIGKPVGHSKSPILYNEAFKSVGFNGVFVHLLVDDIAKFFQTYSSNDF-
CcSDH1        RQMGPDTKVFGIIGKPVGHSKSPILYNEAFKSVGFNGVFVHLLVDDIAKFFQTYSSNDF-
```

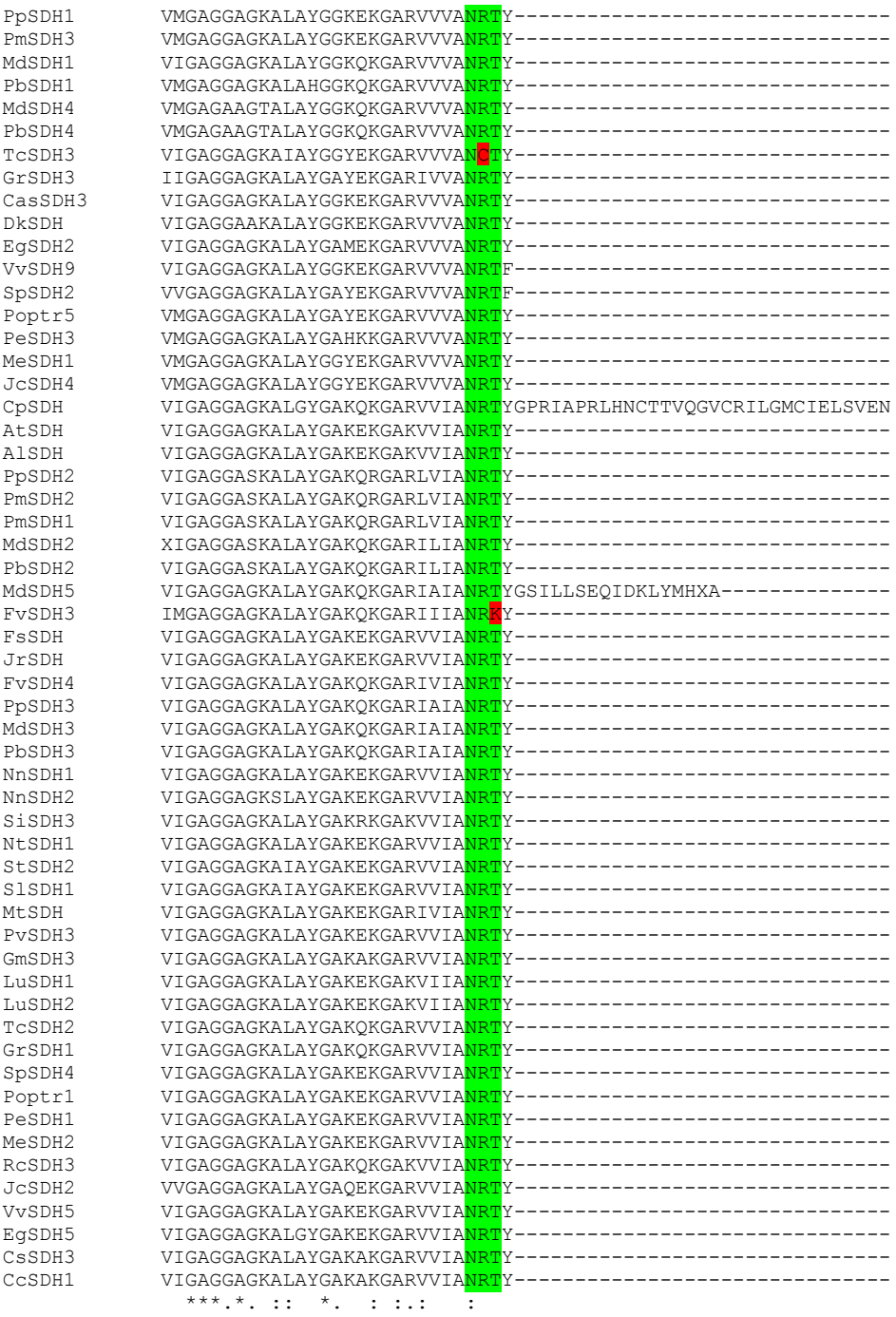

StSDH1 AGFS----------VGIPHKEAAVRCCDEVDPLAKSIGAVNTIIRRPSDGKLIGYNI
SlSDH3 AGFS----------VGIPHKEAAVRCCDEVDPLAKSIGAVNTIIRRPSDGKLIGYNT
SiSDH1 AGFS----------VGIPHKETAVSCCDEVHPLALSIGAVNTIVRRPMDGKLVGYNT
EgSDH1 AGFS----------VGIPHKEAAVACCDEVHPLAKSIGAVNTIVRRPTDGKLIGYNT
GrSDH4 AGFS----------VGIPHKEAAIRCCDEVHPIAKSIGAVNTIVRRPMDGKLIGYNI
CsSDH2 AGFS----------VGIPHKEPAVACCDEVHPLAKSIGAVNTIIRRPIDGKLVGYNT
CcSDH3 AGFS----------VGIPHKEPAVACCDEVHPLAKSIGAVNTIIRRPIDGKLVGYNT
Poptr4 AGFS----------VGIPHKEAAVGCCDEVHPLAKSIGAVNTIVRRPTDGKLVGYNT
SpSDH3 AGFS----------VGIPHKEAAVGCCDEVHPLAKSIGAVNTIVRRPTDGKLVGYNT
JcSDH5 AGFS----------VGIPHKEAAVGCCDEVHPLAKSIGAVNTIVRRPTDGKLIGYNT
RcSDH4 AGFS----------VGIPHKEAAVSCCDEVHPLAKSIGAVNTIVRRPTDGKLVGYNT
MeSDH4 AGFS----------VGIPHKEAAVGCCDEVHPLAKSIGAVNTIVRRPTDGKLVGYNI
SiSDH2 AGFS----------VGIPYKEAVIDFCDEVHPLAQSIGAVNTIVRRPDDGKLVGHNT
PvSDH1 SGFS----------VGIPYKEEVLRFCDELHPLAQSIGAVNTIIRRPGDGKLVGYNT
GmSDH2 SGFS----------VGIPYKEEVLRFCDEVHPLAQSIGAVNTIIRRPRDGKLVGYNT
NtSDH2 AGFS----------VGIPYKEAVVSFCDEVDPLAKSIGAVNTIIQRPCDGKLIGYNT
StSDH3 AGFS----------VGIPYKEAVVSFCDVVDPLAESIGAVNTIIRRPCDGKLIGYNT
SlSDH2 AGFS----------VGIPYKEAVVSFCDEVDPLAESIGAVNTIIRRPCDGKLIGYNT
EgSDH4 AGFS----------VGIPYKEAVIQFCDEVHPLAQNIGAVNTIVRRPSDGKLIGYNI
TcSDH1 AGFS----------VGFPYKEAVVEFCDEVHPLAESIGAVNTIVRRPCDGKLIGYNI
GrSDH2 PGFS----------VGIPYKEAVVEFCDEVHPLAESIGAVNTIIRRPCDGKLIGYNT
CasSDH1 AGFS----------VGFPYKETVTAFCDEVDPLAQSIGAVNTIIRRHSDGKLVGYNT
VvSDH7 AGFS----------VGIPYKEAVTGFCDELHPLAQSIGAVNTIMRRPSDGKLIGYNT
CsSDH1 AGFS----------VGFPYKEAVMKFCDEVHPLAQAIAAVNTIIRRPSDGKLIGYNT
CcSDH2 AGFS----------VGFPYKEAVMKFCDEVHPLAQAIAAVNTIIRRPSDGKLIGYNT
SpSDH1 AGYS----------VGFPYKEAVVQFCDEVHPLAKSIGAVNTIIRRPGDGKLIGYNT
Poptr3 AGYS----------VGFPYKEAVVQFCDEVHPLAKCIGAVNTIIRRPCDGKLIGYNI
PeSDH2 AGYS----------VGFPYKEAVVQFCDEVHPLAKCIGAVNTIIRRPCDGKLVGYNT
Poptr2 AGYS----------VGFPYKEAVVQFCDEVHPLAKSIGAVNTIIRKPSDGKLIGYNT
SpSDH5 AGYS----------VGFPYKEAVVQFCDEVHPLAKSIGAVNTIIRKPSDGKLIGYNT
RcSDH2 AGFS----------VGFPYKEAVVEFCDEVHPLAKSIGAVNTIIRRPGDGKLIGHNT
JcSDH1 AGFS----------VGFPYKEAVVEFCDEVHPLAKSIGAVNTIVRRPSDGKLIGYNT
SpSDH6 SGC----------SCTIPHKEVALKSMDEIDPTAKKIGAINNIVRRP-DGTLKAFNT
TcSDH4 AGC----------SCTIPHKEVALKCMDEIDPIAKKIGAINNIVRQP-DGGLIAFNI
CasSDH2 AGF----------SCTIPHKEVALECMDEIDPIAKKIGAINNIVRRP-DGTLKSYNT
VvSDH8 SGC----------SCTIPHKEVAIKCMDTIDPIARKIGAINNIVRKP-DGKLTAFNT
FvSDH1 SGC----------SCTIPHKEAALKCMDEIDSIAKKIGAINNIVRKP-DGRLVAFNT
EgSDH3 SGC----------SCTIPHKEAALKCMDEIDPIAKKIGAINNIVRRP-DGTLTAFNT
MeSDH3 SGC----------SCTIPHKEVALKCMDEIDPIAKKIGAINNIVRRP-DGTLMAYNT
RcSDH1 SGC----------SCTIPHKEVALKCMDEIDPIAKKIGAINNIIRRP-DATLMAYNT
JcSDH3 SGC----------SCTIPHKEVALKCMDEIDPIARKIGAINNIVRRP-DGTLMAYNT
SbSDH2 AGFS----------CTIPHKEAAVRCCDEVDPIARDIGAVNTIVRRP-DGKLVGYNT
ZmSDH2 AGFS----------CTIPHKEAAVRCCDEVDPVARDIGAVNTIVRRP-DGKLVGYNT
SeiSDH2 AGFS----------CTIPHKEAAVRCCDEVDPIARDIGAVNTMVRRP-DGKLVGYNT
PavSDH2 AGFS----------CTIPHKEAAVRCCDEVDPIARDIGAVNTMVRRP-DGKLVGYNT
PhSDH2 AGFS----------CTIPHKEAAVRCCDEVDPIARDIGAVNTMVRRP-DGKLVGYNT
PavSDH4 AGFS----------CTIPHKEAAVRCCDEVDPIARDIGAVNTMVRRP-DGKLVGYNT
OsSDH1 AGFS----------CTIPHKEAAVRCCDEVDPIAKDIGAVNTIIRKP-NGKLVGYNT
BdSDH2 AGFS----------CTIPHKEAAVRCCDEVDPIARDIGAVNTIVRRS-DGKLVGYNI
TaSDH1 AGFS----------CTIPHKEAAVRCCDEVDPIARDIGAVNTIIRKP-DGKLVGYNT
TaSDH2 AGFS----------CTIPHKEAAVRCCDEVDPIARDIGAVNTIIRKP-DGKLVGYNT
ZmSDH1 AGFS----------CTMPHKETALRCCDDVDPIARDIGAINTIIRRP-DGKLVGYNT
SbSDH1 AGFS----------CTMPHKETALRCCDDVDPIARDIGAINTIIRRP-DGKLVGYNT
SeiSDH1 AGFS----------CTMPHKETAIRCCDDLDPIARDIGAINTIVRRP-DGKLVGYNT
PavSDH3 ----------QETAVRCCDDLDPIARDIGAINTIVRRP-DGKLVGYNT
PhSDH1 AGFS----------CTMPHKETAVRCCDDLDPIARDIGAINTIIRRP-DGKLVGYNI
PavSDH1 AGFS----------CTMPHKEIAVRCCDYLDPIARDIGAINTIVRRP-DGKLVGYNI
SeiSDH3 SGFS----------CSLPFKVDAVQCCDEHDPVAKSIGAINTIIRRP-DGKLVGYNT
PhSDH3 SGFS----------CSLPFKVDAVQCCHEHDPVAKSIGAINTIIRRP-DGKLVGYNT
SbSDH3 SGFS----------CSLPFKVDAVQCCDEHDPVAKSIGAINTIIRRP-DGKLVGYNT
ZmSDH3 SGFS----------CSLPFKVDAVQCCDEHDPVAKSIGAINTIIRRP-DGKLVGYNT
TaSDH3 SGFS----------CSLPFKVDAVHCCHEHDAVAKSIGAINTIVRKS-DGKLVGYNT
OsSDH2 SGFS----------CSLPFKVDAVQCCHEHDPVAKSIGAINTIIRRP-DGKLVGYNT
BdSDH1 SGFS----------CSLPFKVDAVHCCHEHDTVAKSIGAINTIIRKP-DGKLVGYNT
MaSDH3 VGFS----------CTMPHKEIAVRCCDAVDPIAKSIGAVNTIIRRPSDGKFVGHNT
MaSDH4 VGFS----------CTMPHKEIAVRCCDAVDPIAKSIGAVNTIIRRPSDGKFVGHNT
ElgSDH2 AGFS----------CTLPYKETAVTCCDEIDPIAKSIGAINTIIRRPADGKLVGYNT
ElgSDH1 AGFS----------CTIPHKEAAVRCCDEVDPIAKSIGAVNTIIRRPADGKLVGYNT
PdSDH1 ARFS----------CGIPHKEAAVTCCDEIDPIAKSIGAVNTIIRRPADGKLIGYNT
PdSDH2 AGFS----------CTIPHKEAAVTCCDEIDPIAKSIGAVNTIVRRPADGKLVGHNT
MaSDH1 AGFS----------CTIPHKEAAVRCCDEVDPIAKSIGAVNTIVKWPTDGKLIGYNI
MaSDH2 AGFS----------CTIPHKEAAVGCCDKVDPIAKSIGAVNTIIRRPSDGKLIGYNT
PvSDH2 VGYS----------YTIPHKVDGLRCCDEVDPIAKAIGAISCMIKRPSDGKLIGYNL
GmSDH1 VGYS----------YTIPHKQNGLRCCDEVDPIAKAIGAISCMIKRPNDGRLIGYNF
FvSDH2 VGYS----------YTIPHKEAGFKCCDEIDPNALAIGAISCMIRNPTDGKLKGYNV


PpSDH1 VGYS----------FTIPHKESGLKCCDEIDPKAKEIGAVNCMIRRPTDGKLIGYNV
PmSDH3 VGYS----------FTIPHKEAGLKCCDEIDPKAKEIGAVNCMIRRTTDGKLIGYNV
MdSDH1 VGYS----------YTIPHDDAGIKCCDEIDPNAKEIGAISCMIRNPADGKLKGYNM
PbSDH1 VGYS----------YTIPHNDAGIKCCDEIDPTAKEIGAISCMIRNPADGKLKGYNV
MdSDH4 AGYS----------YTIPHKEAGLKCCDEIDPHAKEIGAINCMIRNPADGKLKGYNV
PbSDH4 AGYS----------YTIPHKEAGLKCCDEIDPHAKEIGAINCMIRNPADGKLKGYNV
TcSDH3 AGYS----------SYTIPHKEAGLKCCDEVDPIAESIGAISCMIRRPTDGKLFGYNI
GrSDH3 AGYS----------YTIPHKEAGLKCCDEVDPIAKAIGAISCMIKRSTEGKLIGYNI
CasSDH3 VGYS----------YTIPHKEAGPKCCDEVDPIAKAIGAISCMVRKPSDGKLIGYNV
DkSDH TGYS----------YTIPHKMAGLECCDEVDPIAKAIGAISCMIKKPSDGKLIGYNV
EgSDH2 VGYS----------YTIPHKEDGLRCCDEVDPIAKAIGAISCMIRRPTDGKMIGYNV
VvSDH9 VGYS----------YTIPHKEDGLRCCDEIDPIAQAIGAISCMIRRPADGKLMGYNV
SpSDH2 AGYS----------FTIPHKEDGLKCCNEVDPIAKEIGAISCMIRRPVDGKLKGYNV
Poptr5 AGYS----------YTIPHKEDGLKCCDEVDPIAKEIGAISCMIRRPDDGKLKGYNV
PeSDH3 AGYS----------YTIPHKEDGLKCCDEVDPIAKEIGAISCMIRRPGDGKLEGYNV
MeSDH1 VGYS----------YTIPHKEAGLKCCDEVDPIAQAIGAISCMIRKPTDGKLMGYNV
JcSDH4 VGYS----------YTIPHKEDGLKCCDEIDPIAQAIGAISCMIRRSTDGKLMGYNV
CpSDH AGF----------SCTIPHKEAAMKSCDEVDPVAKSIGAVNCIIRRYNDGKLLGYNT
AtSDH AGF----------SCTIPHKEAALQCCDEVDPLAKSIGAVNTILRRKSDGKLLGYNT
AlSDH AGF----------SCTIPHKEAALQCCDEVDPLAKSIGAVNTILRRKSDGKLLGYNI
PpSDH2 AGF----------SVGIPHKEAALMCCDEVDPVAKSIGAINCIIRRPADGKLFGFNT
PmSDH2 AGF----------SVGIPHKEAALKCCDEVDPVAKSIGAINCIIRRPADGKLFGFNT
PmSDH1 AGF----------SVGIPHKEAALKCCDEVDPVAKSIGAINCIIRRPADGKLFGFNT
MdSDH2 AGF----------SVGIPHKEAAVKCCDEVDPVAKSIGAINCIIRXPTDGKLYGFNT
PbSDH2 AGF----------SVGIPHKEAAVKCCDEVDPVAKSIGAINCIIRRPTDGKLYGFNT
MdSDH5 AGF----------SVTIPHKEAALKCCDEVDPVAKAIGAINCIIRRPTDGKLFGFNT
FvSDH3 SGF----------SVGIPHKEAALKCCDEVDPIAKSIGAVNCIVRRPTDGKLFGLNT
FsSDH AGF----------SITIPHKEAALKCCDEVDPVAKSIGAVNCIVRRSTDGKIFGYNT
JrSDH AGF----------SCTIPHKEAALKCCDEVDPVAKSIGAVNCIIRRPTDGKLVGYNT
FvSDH4 AGF----------SVTIPHKEAALKCCDEVDPVAKSIGAINCIIRRPNDGKLYGLNT
PpSDH3 AGF----------SITIPHKEAALKCCDEVDPVAKSIGAINCIIRRPTDGKLFGFNT
MdSDH3 XGF----------SXTIPHKEAALECCDEVDPVAKAIGAINCIIRRPTDGKLFGFNT
PbSDH3 AGF----------SVTIPHKEAALECCDEVDPVAKAIGAINCIIRRPTDGKLFGFNT
NnSDH1 FGF----------SCTIPHKEDALKRCDEVHPVAKSIGAVNTIIRRSNDGKLSGYNI
NnSDH2 VGFRFNPDKWWRINQSNCTIPHKEAALERCDEVDPIAKSIGAVNMIIRNPSDGKLSGYNC
SiSDH3 SGF----------SITIPHKEVALALCDEVDPTAKAIGAVNCVIRRPTDGHLFGCNT
NtSDH1 AGS----------AVTIPHKEAIVDCCDELNPTAKVIGAVNCVVSRL-DGKLFGCNT
StSDH2 AGF----------SCTIPHKEAALDCCAEIDPTAKAIGAVNCIIRRP-DGKLFGCNT
SlSDH1 AGF----------SCTIPHKEAALDCCAEIDPTAKAIGAVNCIIRRP-DGKLFGCNT
MtSDH VGF----------SVTIPHKEAALKCCDEVDPVAKSIGVVNCVVRRPTDGKLIGYNT
PvSDH3 VGF----------SVTIPHKEAALECCDEVDPVAKSIGAVNCIIRRPTDGKLIGYNT
GmSDH3 VGF----------SVTIPHKETALKCCDEVDPVAKSIGAVNCIVRRPTDGKLIGYNI
LuSDH1 AGF----------SCTIPHKEAAVKCCDEVDPVAKSIGAVNCIVRRESDGKLFGFNT
LuSDH2 AGF----------SCTIPHKEVAVKCCDEVDPVAKSIGAVNCIVRRESDGKLFGFNT
TcSDH2 AGFS----------SCTIPHKEAAVKCCDEVDPIAKSIGAVNCIIRRQSDGKLFGYNT
GrSDH1 AGF----------SCTIPHKETAVKCCDEVDPVAKSIGAVNCIIRRQSDGKLFGCNT
SpSDH4 AGF----------SCTIPHKEDAVKCCDEVHPVAKSIGAANCIIRRQNDGKLFGYNT
Poptr1 AGF----------SCTIPHKEDAAKCCDEVHPVAKSIGAVNCIIRRQNDGKLFGYNT
PeSDH1 AGF----------SCTIPHKEDAVKCCDEVHPVAKSIGAVNCIIRRQNDGKLFGYNI
MeSDH2 ----------CTIPHKEAALKCCDEVDPVAKSIGAVNCIIRRQSDGKLFGYNT
RcSDH3 AGF----------SCTIPHKEAAVDCCDGIDSVARSIGAVNCIIRRQRDGKLFGYNT
JcSDH2 AGF----------SCTIPHKEAAVKCCDEVDPVAKSIGAVNCIIRRQSDGKLFGYNT
VvSDH5 AGF----------SCTIPHKEAALKCCDEVSPVAKSIGAVNCIIRRPSDGKLFGYNT
EgSDH5 AGF----------SCTIPHKEAAVKCCDEVDPVAKSIGAVNCIIRRQSDAKLFGYNT
CsSDH3 AGF----------SCTIPHKEAAVKCCDEVDTVAKSIGAVNCIIRRQSDGKLFGYNT
CcSDH1 AGF----------SCTIPHKEAAVKCCDEVDPVAKSIGAVNCIIRRQSDGKLFGYNT


StSDH1 DCEACVTAIEDALRERQK----------TNG--CASNVSPIAGKLFV
SlSDH3 DCEACVTAIEDALRERQK----------TNG--HASNVSPIAGKLFV
SiSDH1 DCEACITAIEDAFRERFV----------RNG--EALHVSPIAGKLFV
EgSDH1 DCEASITAIEDALRERHAAN----------GEA--RAMDASPIAGKAFV
GrSDH4 DCEAAISAIEDALTDRKVSS----------KGV--PNSTSSLISGRTFV
CsSDH2 DCESAISAIEDALRERQ--G----------ING--VASHTSPIAGKIFV
CcSDH3 DCESAISAIEDALRERQ--G----------ING--VASHTSPIAGKIFV
Poptr4 DCDASISAIEDALTERRI----------TQK--GVLEASPLSGKTFV
SpSDH3 DCDASISAIEDALTERRI----------TQK--GVLEASPLAGKTFV
JcSDH5 DCEAAVTAIEDALRERSMGS----------HNR--GASDGSPLAGKTFV
RcSDH4 DCEASITAIEDALRERR--P----------TNR--GALDTSPLAGKTFV
MeSDH4 DCEASISAIEDALRERQ--A----------TNG--GGSDASPLAGKTFV
SiSDH2 DCEAGITAIEDALKECTV----------L----EPLIPSPLTRKLFV
PvSDH1 DCEAAIAAIEDALIEHGC----------NDG--GASLGSPLAGRLFV
GmSDH2 DCEAAITAIEDALIEHGC----------NDG--EASLGSPLAGRLFV
NtSDH2 DCEASITAIEDALKVNGL----------TNG--AAFLPSPLAGKLFV
StSDH3 DCEASITAIEDALKVNGP----------TNG--AAFLPCSLARKMFV
SlSDH2 DCEASITAIEDALKAN----------G--EALVPCSLARKMFV
EgSDH4 DCEASVTAMEDALQECRC----------ING--EKSLVSPLAGKEFV
TcSDH1 DCEAAITAIEDALKEKRC----------SSN--EASSGTPLSGKLFV
GrSDH2 DCEAAIISIEDALK----------GTLISGKLFV
CasSDH1 DCEASITAIEDALKVWGC----------TNG--EVSLPSPLTGKMFV
VvSDH7 DCEASITAIEDALRERGL----------PNG--EAPLNSPLTGKQFV
CsSDH1 DCEASITAIEDAIKERGY----------KNG--TASFGSPLAGRMFV
CcSDH2 DCEASITAIEDAIKERGY----------KNG--TASFGSPLAGRMFV
SpSDH1 DCDGSITAIEDALRDQKY----------VNG--R-SLNSPLAGKQFV
Poptr3 DCEGSITAIEDALRDQKY----------VNG--R-SLNSPLAGKQFV
PeSDH2 DCEGSITAIEDALRDQKY----------VNG--R-SLNSPLAGKQFV
Poptr2 DCEGSIASIEDALKDQRY----------ING--A-SLNSPLAGKQFV
SpSDH5 DCEGSIAAIEDALKEQKS----------ING--A-SLNSPLAGKQFV
RcSDH2 DCEAAITAIEDALKEQGY----------MDG--RTSSNSPLTGRQFV
JcSDH1 DCEASITAIEDALKEQGY----------ING--RTSFSSSLAGRQFV
SpSDH6 DYIGAISAIEDGLR----------AGASPLSGKLFV
TcSDH4 NYIGAISAIEDGLRGLNV----------VTVTSSAGSPLAGKLFV
CasSDH2 DYIGAISAIEDGLRESNG----------SS--PATSSPLAGKLFV
VvSDH8 DYIGAIEAIEDGLRESNG----------SS--PAVGSPLAGKLFV
FvSDH1 DYIGSISGIEDELRGMNG----------AI--PAGKSPLAGKLFV
EgSDH3 DYIGAISAIEDGLRGLNV----------V---SPGASPLAGKLFV
MeSDH3 DYIGAISAIEDGLRELNG----------KV--PAGTSPLAGKLFV
RcSDH1 DYIGAIDAIEDGLRELNG----------AV--PAGTSPLAGKLFV
JcSDH3 DYIGAISAIEDGLRELNG----------AV--PAGTSPLKGKLFV
SbSDH2 DYVGAISAIEDGIRAS----------E--QTDSDTSPLAGRLFV
ZmSDH2 DYVGAISAIEDGIKAS----------ESEPTDPDKSPLAGRLFV
SeiSDH2 DYVGAISAIEDGIRAS----------Q----DPSTSPLAGRLFV
PavSDH2 DYVGAISAIEDGIRAS----------Q--QTDPTTSPLAGRLFI
PhSDH2 DYVGAISAIEDGIRAS----------Q--QTDPTTSPLAGRLFV
PavSDH4 DYVGAISAIEDGIRAC----------Q--LTDPTTSPLAGRLFV
OsSDH1 DYVGAISAIEDGIRAS----------Q--PTDTTTSPLAGRLFV
BdSDH2 DYVGAISAIEDGIRAS----------Q--STHSTASPLAGRLFV
TaSDH1 DYVGAISAIEDGIRAT----------Q--PTHSTTSPLAGRLFV
TaSDH2 DYVGAISAIEDGIRAT----------Q--QTHSTVSPLAGRLFV
ZmSDH1 DYSGAISAIEDGIRAS----------Q--PTDSITSPLAGRLFV
SbSDH1 DYSGAISAIEDGIRAS----------Q--STDSTTSPLAGRLFV
SeiSDH1 DYVGAISAIEDAIRAS----------H--PVDPTTSPLARRLFV
PavSDH3 DYVGAIAAIEDAIRAS----------Q--PADPTTSLLARRLFV
PhSDH1 DYVGAIAAIEDAIRAS----------Q--PTDPTTSPLARRLFV
PavSDH1 DYVGAIAAIEDAIRAS----------Q--PTDPTTSPLAGRPFV
SeiSDH3 DYIGAISAIEDGIG----------GPGSKDAAISLLAGRLIV
PhSDH3 DYIGAISAIEDGIG----------GPGSKVAAISPLAGRLIV
SbSDH3 DYIGAISAIEDGIG----------GPGSKDAASSPLAGRLVV
ZmSDH3 DYIGAISAIEDGIG----------GPGSKDAVISPLAGRLIV
TaSDH3 DYIGAISAIEDGIGAHDL----------CSTGLGSKDAAMSPLSGRLIV
OsSDH2 DYIGAISAIEDGIG----------GPGSKDAAISPLAGRLVV
BdSDH1 DYIGAISAIEDGIG----------GPGSKDAAVSPLAGRLIV
MaSDH3 DYFGAISAIEEELRGALRYTYMETAARKDHILFSLFHGTGSQGQGKEEETASPLAGRVFV
MaSDH4 DYFGAISAIEEELRG----------SQGQGKEEETASPLAGRVFV
ElgSDH2 DYIGAISAIEDGIRGS----------WA---KDAVVSPLAGKLFV
ElgSDH1 DYIGAISAIEDGIRGS----------RA---KDAAVSPLAGKLFV
PdSDH1 DYIAAISAIEDGIRGT----------WA---KDAAVSPLAGKLFV
PdSDH2 DYIGAISAIEDGIRGS----------RA---KDAAVSPLAGKLFV
MaSDH1 DYVGAISAIEDGIRGS----------KGVG-KDETVSPLAGKVFV
MaSDH2 DYVGAISAIEDGIRGS----------QGVG-K-ETVSPLAGKVFV
PvSDH2 DYLGAIAAIEERLH--DS----------NG-RSMSG-SPLYGKLFV
GmSDH1 DYLGAIAAIEERLHLQDS----------NG-RSISGCSPLYGKLFV
FvSDH2 DYLGAIAAIEEGLRGLGL----------NG-SNNGSGSPLAGRLFV

```
PpSDH1      DYLGAIAAIEEGLQA--L----------------------NG-SSNTSGSLLAGKLFV
PmSDH3      DYLGAIAAIEEGLQA--L----------------------NG-SSNTSGSLLAGKLFV
MdSDH1      DYLGAIAAIEEGLQA--L----------------------NG-SSNGSGSPLAGKLFV
PbSDH1      DYLGAIAAIEEGLQA--I----------------------NG-SSNGSGSQLAGKLFV
MdSDH4      DYLGAIAAIEEGLRA--L----------------------NG-SSNGSGSPLAGKLFV
PbSDH4      DYLGAIAAIEEGLRA--L----------------------NG-SSNGSGSPLAGKLFV
TcSDH3      DYLGAIAAIEEG--LRAT----------------------SG-ATPASGSPLAGQIFV
GrSDH3      DYLGGIAAIEEE--LRVS----------------------NG-ATLASGSPLAGKTFV
CasSDH3     DYLGAIGAIEEA--LGGS----------------------NG-SSSGAVSPLASKLFV
DkSDH       DYLGAIAAIEEG--LGGS----------------------SS-ASNGSVSPLAGRLFV
EgSDH2      DYLGAIAAIEEA--LRAS----------------------NG-ASSTTTSPLAGKLFV
VvSDH9      DYLGAIAAIEEG--LRAS----------------------NG-T-TSVGSPLAGKLFV
SpSDH2      DYLGAIKTIEEALALGAS----------------------NG-A-PASVSPLAGKLFV
Poptr5      DYLGAIAAIEEA--LGAS----------------------NG-A-PASVSPLAGKLFV
PeSDH3      DYLGAIAAIEEA--LGAS----------------------NG-A-PASVSPLAGKLFV
MeSDH1      DYLGAIGAIEEA--LQGS----------------------NG-S-PASGSPLAGKLFV
JcSDH4      DYLGAIGAIEEA--LRAS----------------------NG-S-STSGSPLAGKLFV
CpSDH       DYVGAISAIEDRLQGNAN----------------------------SAAGSPLAGKLFV
AtSDH       DCIGSISAIEDGLRSSGD----------------------PSSV-PSSSSPLASKTVV
AlSDH       DCIGSISAIEDGLRRSGD----------------------PSSG-PSSSSPLAGKTVV
PpSDH2      DYVGAISAIEEGLKGSHN----------------------NSGKNTTSGSALAGRLFV
PmSDH2      DYVGAISAIEEGLKGSHN----------------------NSGKNTTSGSALAGRLFV
PmSDH1      DYVGAISAIEEGLKGSHN----------------------NSGKNTTSGSALAGRLFV
MdSDH2      DYVGAISAIEEGLKGSPN----------------------AG---SVTGSPLAGRLFV
PbSDH2      DYVGAISAIEEGLKGSPN----------------------AG---SVTGSPLAGRLFV
MdSDH5      DYVGAISAIEDGLKGSHN----------------------NG---SLXVSPLAGRLFV
FvSDH3      DYFGSISAIEDGLRGSHD----------------------KS---NIIGSPLAGRLFV
FsSDH       DYVGAISAIEDGLRGSKN----------------------IS---YAADSPLAGKLFV
JrSDH       DYVGAISAIEDGLRGSHN----------------------SS---NTADSPLAGKLFV
FvSDH4      DYVGAISAIEDGLQGSLN----------------------GS---HVTGSPLAGRLFV
PpSDH3      DYVGAISAIEDGLTGSHK----------------------SG---SITGSPLAGRLFV
MdSDH3      DYVGAISAIEDGLKGSHN----------------------NG---SLTGSPLAGRLFV
PbSDH3      DYVGAISAIEDGLKGSHN----------------------SG---SLTGSPLAGRLFV
NnSDH1      DYAGAISAIEDGLRGLHD----------------------MH---NTIASPLAGKLFI
NnSDH2      DYIGAISAIEHGLRGSHD----------------------IH---NTVSSPLAGKLFV
SiSDH3      DYIGAISAIEDGLLGSHH----------------------GP---TGAGSPLAGKLFV
NtSDH1      DYVGAISAIEEALQGSQ-----------------------P---SMSGSPLAGKLFV
StSDH2      DYIGAISAIEEGLQGSQ-----------------------P---SMSGSPLAGKLFV
SlSDH1      DYIGAISAIEEGLQGSQ-----------------------P---SISGSPLAGKLFV
MtSDH       DYVGAISAIEDGLRGKHN----------------------DS---GTAVSPLVGKLFV
PvSDH3      DYVGAISAIEDGLRGKHN----------------------GG---STNISPLAGKLFV
GmSDH3      DYVGAITAIENGLRGKHN----------------------GS---STTISPLAGKLFV
LuSDH1      DYVGAISAIEDGLQASRN----------------------GS---SQSGSPLAGKAFV
LuSDH2      DYVGAISAIDDGLQASRN----------------------GS---SQSGSPLAGKVFV
TcSDH2      DYVGAISAIEDGLPARLN----------------------TS---NTAGSPLAGKRFV
GrSDH1      DYIGAISAIEDGLQAGYN----------------------MS---SAAGSPLAGKLFV
SpSDH4      DYVGAISAIEDGLRASQN----------------------VS---NTVGSPLAGKLFI
Poptr1      DYVGAISAIEEGLRASQN----------------------VS---NTVGSPLAGKLFV
PeSDH1      DYVGAISAIEEGLRASQN----------------------VS---DTVGSPLSGKLFV
MeSDH2      DYVGAISAIEEGLQGSQN----------------------KI---GVTVSPLAGKLFV
RcSDH3      DYDGAITAIEEGLRVSQN----------------------IS---GAADSPLAGKLFV
JcSDH2      DYVGAISAIEEGLRVSQN----------------------IS---STAGSPLAGKLFV
VvSDH5      DYVGAISAIEDGLRDLHK----------------------IS---STSGSPLAGKLFV
EgSDH5      DYVGAISAIEDGLRGSQN----------------------GN---SAGASPLNGKLFV
CsSDH3      DYVGAISAIEDGLRGRLN----------------------VS---GGVSSALAGKLFV
CcSDH1      DYVGAISAIEDGLRGRLN----------------------VS---GGVSSALAGKLFV
            :  . :  ::.                                      :  : .:
```


StSDH1 LVGAGGAGRAIAFGAKSRGARVVIFNRKY
SlSDH3 LVGAGGAGRAIAFGAKSRGARVVIFNRKY
SiSDH1 LVGAGGAGRAIAFGARSRGARVVIFNRNF
EgSDH1 LVGAGGAGRALAFGAKSRGARVFVFNRNF
GrSDH4 LIGAGGAGRALAFGAKYKGARIVIFNRDY
CsSDH2 LVGAGGAGRALAFGAKSRGARVIIFNRNY
CcSDH3 LVGAGGAGRALAFGAKSRGARVIIFNRNY
Poptr4 LIGAGGAGRALAFGAKSRGARVIIFNRNY
SpSDH3 LIGAGGAGRALAFGAKSRGARVIIFNRDY
JcSDH5 LVGAGGAGRALAFGAKSRGAQIVIFNRNY
RcSDH4 LIGAGGAGRALAFGASSRGARVVVFNRNF
MeSDH4 LVGAGGAGRALAFGAKSRGAHVIVFNRNY
SiSDH2 LVGAGGAGRALAFGAKTRGARVVIFDIDF
PvSDH1 LVGAGGAGKALAFGAKSRGARLVIFDIDF
GmSDH2 LVGAGGAGKALAFGAKSRGARLVIFDIDF
NtSDH2 LVGAGGAGRALAFGAKSRRAEIVIFDIDF
StSDH3 LVGAGGAGRALAFGAKSRGARIVIFDIDF
SlSDH2 LVGAGGAGRALAFGAKSRGARIVIFDIDF
EgSDH4 LVGAGGAGRALAFGAKTRGARIIIFDIDF
TcSDH1 LVGAGGAGRALAFGAKSRGSRIVIFDIDF
GrSDH2 LVGAGGAGRALAFGAKSRGARILIFDIDF
CasSDH1 LVGAGGAGRALAFGAKSRGARVVIFDIDF
VvSDH7 LVGAGGAGRALAFGARSRGAQLVIFDLDF
CsSDH1 LAGAGGAGRALAFGAKSRGARVVIFDIDF
CcSDH2 LAGAGGAGRALAFGAKSRGARVVIFDIDF
SpSDH1 VVGAGGAGRAIAVGARSRGARLIVFDIDL
Poptr3 VVGAGGAGRAIAVGAKSRGARLIIFDIDL
PeSDH2 VVGAGGAGRAIAVGAKSRGARVIIFDIDL
Poptr2 VVGAGGAGRAIAVGAKSRGARVIIFDIDL
SpSDH5 VAGAGGAGRAIAVGAKSRGARVIIFDIDL
RcSDH2 LVGAGGAGRALAFGAKSRGARIVIFDIDL
JcSDH1 LVGAGGAGRALAFGAKSRGARIIIFDIDF
SpSDH6 VLGAGGAGKSLAYGAAQKGARVVVANRTF
TcSDH4 VLGAGGAGKSLAYGAAQKGARIIIANRTF
CasSDH2 VLGAGGAGKSLAYGAQQKGARVVVANRTF
VvSDH8 VLGAGGAGKSLAYGAKEKGARVVVANRTF
FvSDH1 VLGAGGAGKSLVYGAAQKGARVVCANRTY
EgSDH3 VLGAGGAGKSLAYGAAQKGARVVVANRTF
MeSDH3 VLGAGGAGKSLAYGGAQKGARVVVANRSY
RcSDH1 VLGAGGAGKSLAYGAAQKGARVVVANRTF
JcSDH3 VLGAGGAGKSLAYGAAQKGARVVVANRTY
SbSDH2 VIGAGGAGKALAYGAKEKGARVVIANRTF
ZmSDH2 VIGAGGAGKALAYGAKEKGARVVIANRTF
SeiSDH2 VIGAGGAGKALAYGAKEKGARVVIANRTF
PavSDH2 VIGAGGAGKALAYGAKEKGARVVIANRTF
PhSDH2 VIGAGGAGKALAYGAKEKGARVVIANRTF
PavSDH4 VIGAGGAGKALAYGAKEKGARVVIANRTF
OsSDH1 VIGAGGAGKALAYGAKEKGARVVIANRTF
BdSDH2 VIGAGGAGKALAYGAKEKGARVVIANRTF
TaSDH1 VIGAGGAGKALAYGAKEKGARVVIANRTF
TaSDH2 VIGAGGAGKALAYGAKEKGARVVIANRTF
ZmSDH1 VIGAGGAGKALAYGAKEKGARVVIANRTF
SbSDH1 VIGAGGAGKALAYGAKEKGARVVIANRTF
SeiSDH1 VIGAGGAAKAVAYGAKEKGARVVIANRTF
PavSDH3 VIGAGGAAKAVAYGAKEKGARVVIANRTF
PhSDH1 VMGAGGAAKAVAYGAKEKGARVVIANRTF
PavSDH1 VIGAGGAAKAVAYGAKEKGAKVVIANRTF
SeiSDH3 VVGAGGAGKAIAYGAKEKGARVVVANRTY
PhSDH3 VVGAGGAGKAIAYGAKEKGARVVVANRTY
SbSDH3 VVGAGGAGKAIAYGAKEKGARVVIANRTY
ZmSDH3 VVGAGGAGKAIAYGAKEKGARVVIANRTY
TaSDH3 VVGAGGAAKAIAYGAKKKGARVVVANRTY
OsSDH2 VVGAGGAGKAIAYGAKEKGARVVVANRTY
BdSDH1 VVGAGGAGKAIAYGAKKKGARVVVANRTY
MaSDH3 VMGAGGAGKAIAYGAKEKGARVVIANRTY
MaSDH4 VMGAGGAGKAIAYGAKEKGARVVIANRTY
ElgSDH2 VIGAGGAGKALAYGATEKGARVAIANRTY
ElgSDH1 VIGAGGAGKALAYGAKEKGARIVIANRTY
PdSDH1 VIGAGGAGKALAYGAKEKGARVVIANHTY
PdSDH2 VIGAGGAGKALAYGAKEKGARVVIANRTY
MaSDH1 VIGAGGAGKALAYGAKEKGAKVVIANRTY
MaSDH2 IIGAGGAGKALAYGAKEKGAKVVIANRTY
PvSDH2 VIGAGGAGKALAYGGKQKGARIVVANRTY
GmSDH1 VMGAGGAGKALAYGGKEKGARVVVANRTY
FvSDH2 VMGAGGAGKALAYGGKQKGARVVVANRSF


PpSDH1 VMGAGGAGKALAYGGKEKGARVVVANRTY
PmSDH3 VMGAGGAGKALAYGGKEKGARVVVANRTY
MdSDH1 VIGAGGAGKALAYGGKQKGARVVVANRTY
PbSDH1 VMGAGGAGKALAHGGKQKGARVVVANRTY
MdSDH4 VMGAGAAGTALAYGGKQKGARVVVANRTY
PbSDH4 VMGAGAAGTALAYGGKQKGARVVVANRTY
TcSDH3 VIGAGGAGKAIAYGGYEKGARVVVANCTY
GrSDH3 IIGAGGAGKALAYGAYEKGARIVVANRTY
CasSDH3 VIGAGGAGKALAYGGKEKGARVVVANRTY
DkSDH VIGAGGAAKALAYGGKEKGARVVVANRTY
EgSDH2 VIGAGGAGKALAYGAMEKGARVVVANRTY
VvSDH9 VIGAGGAGKALAYGGKEKGARVVVANRTF
SpSDH2 VVGAGGAGKALAYGAYEKGARVVVANRTF
Poptr5 VMGAGGAGKALAYGAYEKGARVVVANRTY
PeSDH3 VMGAGGAGKALAYGAHKKGARVVVANRTY
MeSDH1 VMGAGGAGKALAYGGYEKGARVVVANRTY
JcSDH4 VMGAGGAGKALAYGGYEKGARVVVANRTY
CpSDH VIGAGGAGKALGYGAKQKGARVVIANRTYGPRIAPRLHNCTTVQGVCRILGMCIELSVEN
AtSDH VIGAGGAGKALAYGAKEKGAKVVIANRTY
AlSDH VIGAGGAGKALAYGAKEKGAKVVIANRTY
PpSDH2 VIGAGGASKALAYGAKQRGARLVIANRTY
PmSDH2 VIGAGGASKALAYGAKQRGARLVIANRTY
PmSDH1 VIGAGGASKALAYGAKQRGARLVIANRTY
MdSDH2 XIGAGGASKALAYGAKQKGARILIANRTY
PbSDH2 VIGAGGASKALAYGAKQKGARILIANRTY
MdSDH5 VIGAGGAGKALAYGAKQKGARIAIANRTYGSILLSEQIDKLYMHXA
FvSDH3 IMGAGGAGKALAYGAKQKGARIIIANRKY
FsSDH VIGAGGAGKALAYGAKEKGARVVIANRTY
JrSDH VIGAGGAGKALAYGAKEKGARVVIANRTY
FvSDH4 VIGAGGAGKALAYGAKQKGARIVIANRTY
PpSDH3 VIGAGGAGKALAYGAKQKGARIAIANRTY
MdSDH3 VIGAGGAGKALAYGAKQKGARIAIANRTY
PbSDH3 VIGAGGAGKALAYGAKQKGARIAIANRTY
NnSDH1 VIGAGGAGKALAYGAKEKGARVVIANRTY
NnSDH2 VIGAGGAGKSLAYGAKEKGARVVIANRTY
SiSDH3 VIGAGGAGKALAYGAKRKGAKVVIANRTY
NtSDH1 VIGAGGAGKALAYGAKEKGARVVIANRTY
StSDH2 VIGAGGAGKAIAYGAKEKGARVVIANRTY
SlSDH1 VIGAGGAGKAIAYGAKEKGARVVIANRTY
MtSDH VIGAGGAGKALAYGAKEKGARIVIANRTY
PvSDH3 VIGAGGAGKALAYGAKEKGARVVIANRTY
GmSDH3 VIGAGGAGKALAYGAKAKGARVVIANRTY
LuSDH1 VIGAGGAGKALAYGAKEKGAKVIIANRTY
LuSDH2 VIGAGGAGKALAYGAKEKGAKVIIANRTY
TcSDH2 VIGAGGAGKALAYGAKQKGARVVIANRTY
GrSDH1 VIGAGGAGKALAYGAKQKGARVVIANRTY
SpSDH4 VIGAGGAGKALAYGAKEKGARVVIANRTY
Poptr1 VIGAGGAGKALAYGAKEKGARVVIANRTY
PeSDH1 VIGAGGAGKALAYGAKEKGARVVIANRTY
MeSDH2 VIGAGGAGKALAYGAKEKGARVVIANRTY
RcSDH3 VIGAGGAGKALAYGAKQKGAKVVIANRTY
JcSDH2 VVGAGGAGKALAYGAQEKGARVVIANRTY
VvSDH5 VIGAGGAGKALAYGAKEKGARVVIANRTY
EgSDH5 VIGAGGAGKALGYGAKEKGARVVIANRTY
CsSDH3 VIGAGGAGKALAYGAKAKGARVVIANRTY
CcSDH1 VIGAGGAGKALAYGAKAKGARVVIANRTY

```
StSDH1      ---------------------------------ERAKALAAAVSCEALPYEHLNDFCPEKG
SlSDH3      ---------------------------------ERAKALAAAVSCDALPYEHLNDFCPEKG
SiSDH1      ---------------------------------ERAKALAMAVSGEALPYESLNDFRAEKG
EgSDH1      ---------------------------------ERAKALADAVSGEAIRYEHLDTFRPEEG
GrSDH4      ---------------------------------KRAKALADEVSGEALLYEALEMFRPETG
CsSDH2      ---------------------------------ERAKALADAVSGEALHFEYLHEFFPEKG
CcSDH3      ---------------------------------ERAKALADAVSGEALHFEYLHEFFPEKG
Poptr4      ---------------------------------ERARALAKAVSGEALPYESLDRFRPVNG
SpSDH3      ---------------------------------ERARALAKAVSGEALPYESLDRFRPENG
JcSDH5      ---------------------------------ERAKALAHAVSGKALPYESLKSFCPENG
RcSDH4      ---------------------------------ERAKSLANAVSGEALPYKALDRFQREND
MeSDH4      ---------------------------------ERAKSLAHAVSGKALPYESLEKFRAEKV
SiSDH2      ---------------------------------DRAKSLAAAVSGIALPFEDLIGFRPEKD
PvSDH1      ---------------------------------DRAKSLACAVSGEAQPYKELVNFQPEKG
GmSDH2      ---------------------------------DRAKSLACAVFGEAQPFKELVNFQPEKG
NtSDH2      ---------------------------------DRAKALAAAVSGEALPFENLASFQPEKG
StSDH3      ---------------------------------NRAKVLAAAVCGEALPYEKLASFQPEKG
SlSDH2      ---------------------------------DRAKALAAAVSGEALPFEKLASFQPEKG
EgSDH4      ---------------------------------ERAKMLAHAVSGEARPFGDLPYFQPEKG
TcSDH1      ---------------------------------ERAKSLACAVSGEARVFEEVVNFQPEKG
GrSDH2      ---------------------------------ERAKSLASAVLGEARTFEQVINFRPEKG
CasSDH1     ---------------------------------DRAKSLALAVSGEAQPFENLVSFQPEKG
VvSDH7      ---------------------------------DRANSLAHAVSGEVKLYEDVANFQPEKG
CsSDH1      ---------------------------------ERAKSLASDVMGAARPFEDILNFQPEKG
CcSDH2      ---------------------------------ERAKSLASDVMGAARPFEDILNFQPEKG
SpSDH1      ---------------------------------ERAKCLARAVSGEARHFESLAHFQPENG
Poptr3      ---------------------------------ERAKSLARAVSGEAQHFESLAHFQPENG
PeSDH2      ---------------------------------ERAKSLARAVSGEAQHFESLAHFQPENG
Poptr2      ---------------------------------DRAKSLAQVVSGEAQHFDSLAHFQPEKG
SpSDH5      ---------------------------------TRAESLAQVVSGEAKPFESLAHFQPENG
RcSDH2      ---------------------------------ERAKFLADAVSGEAQLFENVVNFEPENG
JcSDH1      ---------------------------------ERTKSLAHAVSGEAQPFKNLVHFRPETD
SpSDH6      ---------------------------------ERAKELADKVGGLAMTLAEAENFHPEDG
TcSDH4      ---------------------------------KRAKELANKDGGQAMPLAEVENFYPEEG
CasSDH2     ---------------------------------ERAKELAEKVGGKALTLEEVNDFHPEEG
VvSDH8      ---------------------------------ERAKDLADKVGGQALTLAEIENFHPEEG
FvSDH1      ---------------------------------ERAKELADKVGGQAMTLEEVENFHPEEG
EgSDH3      ---------------------------------ERAKELADKVGGQAMTLAEVENFHPEEG
MeSDH3      ---------------------------------ERAKELADKVGGQAMALAEVEHFHPEEG
RcSDH1      ---------------------------------ERAKELADKVNGQAMTLDEVQNFHPEEG
JcSDH3      ---------------------------------ERAKELADKVGGQAMTLAEAEHFHPEDG
SbSDH2      ---------------------------------ARAQELGNLIGGPALTLADLENYHPEEG
ZmSDH2      ---------------------------------ARAQELANLIGGPALTLADLENYHPEEG
SeiSDH2     ---------------------------------ARAQELANLIGGPALTLADLENYHPEEG
PavSDH2     ---------------------------------ARAQELANLIGGPALTLADLENYHPEEG
PhSDH2      ---------------------------------ARAQELGNLIGGPALTLADLENYHPEEG
PavSDH4     ---------------------------------ARAQELANLIGGPALTLADLENYHPEEG
OsSDH1      ---------------------------------ARAQELGNLLGAPALTLAELENYHPEEE
BdSDH2      ---------------------------------ARAQELANLIGGPALTLSELENYHPEEG
TaSDH1      ---------------------------------ARAQELGNLIGGPALTLTELENYHPEEG
TaSDH2      ---------------------------------ARAQELGNLIGGPALTLTELENYHPEEG
ZmSDH1      ---------------------------------ARAQELAHIIGGTALTLSELESYHPEKG
SbSDH1      ---------------------------------ARAQELANKIGGTALTLSELENYHPEEG
SeiSDH1     ---------------------------------ARAQELARLIGGTALTMAELESYHPEDG
PavSDH3     ---------------------------------VRAQHLATLIGGTALTLSELENYHPEEG
PhSDH1      ---------------------------------ARAQELANLIGGTALTLSELENYHPEEG
PavSDH1     ---------------------------------ARAQELATLIGGTALTLSELENYHPEEG
SeiSDH3     ---------------------------------EKAVSLANAIGGQALRLADLETFRPEEG
PhSDH3      ---------------------------------EKAVSLANAIGGQALRLADLETFRPEEG
SbSDH3      ---------------------------------EKAVSLANAIGGQALRLEDLETFRPEEG
ZmSDH3      ---------------------------------EKAVRLANAIGGQALRLEDLETFRPEEG
TaSDH3      ---------------------------------EKAVTLANAVGGQALRLADLENFRPEEG
OsSDH2      ---------------------------------EKAVSLAAAVGGHALRLAELETFRPEEG
BdSDH1      ---------------------------------EKAVGLANAVGGQALRLADLENFRPEGG
MaSDH3      ---------------------------------ERARELAKQVGGQALPLSELENFHPEDG
MaSDH4      ---------------------------------ERARELAKQVGGQALPLSELENFHPEDG
ElgSDH2     ---------------------------------ERARELANLIGGHALTLAELERFHPKEG
ElgSDH1     ---------------------------------ERARELANLIGGRALTLAELESFHPEEG
PdSDH1      ---------------------------------ERARELANLIGGLALTLAELKSFHPEGG
PdSDH2      ---------------------------------ERARELANLIGGLALTLAELESFHPEGG
MaSDH1      ---------------------------------ERARELANLIGGHALTLSELEKFHPEDG
MaSDH2      ---------------------------------ERARELANLIGGHALTLAELETFHPEDG
PvSDH2      ---------------------------------AKAKELASRVGGKAITLSELENFHPEQG
GmSDH1      ---------------------------------AKAKELATKVGGEAITLSELESFHPEQG
FvSDH2      ---------------------------------DKAKILADKVGGEAITLAELENFHPEDG
```

```
PpSDH1      --------------------------------DKAKALANKVGGEAITLAELENFHPEDG
PmSDH3      --------------------------------DKAKALANKVGGEAITLAELENFHPEDG
MdSDH1      --------------------------------DKAKALANKVGGEAITLAELENFHPEEG
PbSDH1      --------------------------------DKAKALANKVGGEAITLAELENFHPEEG
MdSDH4      --------------------------------DKAKALANKVGGEAITLAELETFHPEDG
PbSDH4      --------------------------------DKAKALANKVGGEAITLAELETFHPEDG
TcSDH3      --------------------------------DKAKELASKVGGQAITLSELNEFHPEEG
GrSDH3      --------------------------------DKAKELASKVRGQALTLAELNDFCPEEG
CasSDH3     --------------------------------EKAKELASKVDGEAITLAELDDFHPEDG
DkSDH       --------------------------------EKAKDLARKIGGESMPLTELNDFHPEDG
EgSDH2      --------------------------------EKAKELASKVGGQAITLAELENFHPEDG
VvSDH9      --------------------------------EKAKELASKVGGEAMTLAELENFHPEDG
SpSDH2      --------------------------------GKAKELASKVGGQAITLAELKDFHPEEG
Poptr5      --------------------------------GKAKELASKVGGQAIALAKLKDFHPEEG
PeSDH3      --------------------------------GKAKALAGKVGGQAIPLAKLKDFHPEEG
MeSDH1      --------------------------------EKAKELASKVGGEAMTLAELKEFHPEEG
JcSDH4      --------------------------------EKAKELASKVGGQAMTLDELRDFHPEEG
CpSDH       ERKLKTIWQKSVEEAHPKKSQLPGVYKVEGGKGK-------------------------
AtSDH       --------------------------------ERALELAEAIGGKALSLTDLDNYHPEDG
AlSDH       --------------------------------ERALELAEAIGGRALSLTDLDNFHPEDG
PpSDH2      --------------------------------DRARELADIVGGDALSLDDLPNIRPEDG
PmSDH2      --------------------------------DRARELADIVGGDALSLDDLPDIPPEDG
PmSDH1      --------------------------------DRARELADIVGGDALSLDDLPDIPPEDG
MdSDH2      --------------------------------DRARELADTVGGDAVSLADLCNIPPEDG
PbSDH2      --------------------------------DRARELADTVGGDAVSLADLGNIPPEDG
MdSDH5      GRSNLNTWKEGMLFRELLESTLQLSDTVGGGGDRARELADTIGGDALSLADLDNFHPEDG
FvSDH3      --------------------------------DRARKLADEVGGDALPFADLANFHPEDG
FsSDH       --------------------------------ERAKEIADTVGGQALSLADLENFHPEDG
JrSDH       --------------------------------DRARELADTIGGDALSLADLDNFHPEDG
FvSDH4      --------------------------------DRAREIADTIGGEALSISDLDKFHPEDG
PpSDH3      --------------------------------DRARELADTIGGDALSLADLDNFHPEEG
MdSDH3      --------------------------------DRARELADTIGGDALSLADLDNFHPEDG
PbSDH3      --------------------------------DRARELADTIGGDALSLADLDNFHPEDG
NnSDH1      --------------------------------SRAKELADMVGGDALSLDELGDFHPENG
NnSDH2      --------------------------------ERAKQLANVIGGEAIPLDELGDFHPENK
SiSDH3      --------------------------------DRARELAEIVGGQALSLAELGDFHPESG
NtSDH1      --------------------------------ERARELADVVGGQALSLDELSNFHPEND
StSDH2      --------------------------------ERARELAIVIGAEALSLDELSNFHPDND
SlSDH1      --------------------------------ERARELAIVVGAEALSLDELSNFHPEND
MtSDH       --------------------------------DRARELADVIGGDALTLNDLDNYHPEDG
PvSDH3      --------------------------------DRARELADAIGGDALTLSDLDSYHPEDS
GmSDH3      --------------------------------DHARKLAYAIGGDALALADLDNYHPEDG
LuSDH1      --------------------------------GRAKELADMIGGEAISLADLENFHPESG
LuSDH2      --------------------------------GRAKELADIIGGEAISLADLENFHPESD
TcSDH2      --------------------------------ERARELADIIGGDALSLADLACFHPEEG
GrSDH1      --------------------------------ERARELADVIGGDALSLDDLARFHPEDG
SpSDH4      --------------------------------ERAKVLADIIGGDAISLADLENFHPEDG
Poptr1      --------------------------------ERAKVLADIIGGDAITLADLENFHPEDG
PeSDH1      --------------------------------ERAKVLADIIGGAAITLADLENFHPEDG
MeSDH2      --------------------------------ERAKELADIVGGDAISLADLDNFHPEDG
RcSDH3      --------------------------------ERAKGLADAVGGVAIALADLGNFRPESG
JcSDH2      --------------------------------DRARELAEIVGGDAISLADLENFHPEDG
VvSDH5      --------------------------------ARARELADAVGGDALSLADLNNFHPENG
EgSDH5      --------------------------------DRARELAETIGGDALSLADLENFHPEDG
CsSDH3      --------------------------------DRARELAETVGGHALSLADLENFNPEDG
CcSDH1      --------------------------------DRARELAETVGGHALSLADLENFNPEDG
                                             :
```

```
StSDH1        MILANASAVGMQPKSD-QTPISKEALRSYELVFDAVYTPRNTRLLQEATEVGATVVSGVE
SlSDH3        MILANASAVGMQPKSD-QTPISKEALRSYELVFDAVYTPRNTRLLQEATEVGATVVSGVE
SiSDH1        MILANASAVGMEPNTD-QTPVSKVALKSYELVFDAVYTPRNTRLLQEAAEVGATVVSGVE
EgSDH1        MILANASAVGMEPHAD-KSPVSKGVLGAYELVFDAVYTPRNTRLLQEAAQAGAIAVSGME
GrSDH4        AILANASAIGMEPNSH-QSPVSKEALRAYDLVFDVVYTPRNTQLLKEAAEVGVTTVSGVE
CsSDH2        MILANASAIGMEPNSD-QSPVPKEALKAYELVFDAVYTPRNTRLLREAAEVGATVVSGVE
CcSDH3        MILANASAIGMEPNSD-QSPVPKEALKAYELVFDAVYTPRNTRLLREAAEVGATVVSGVE
Poptr4        MILANASAIGMEPNSD-QSPVSKEILKACELVFDAVYTPRNTRLLREAKEVGAVVVSGVE
SpSDH3        MILANASAIGMEPNSD-QSPVSKEALKAYELVFDAVYTPRNTRLLQEAKEAGAVVVSGVE
JcSDH5        MILANASAVGMEPNPD-QTPVSKDALKAYDLVFDAVYTPRNTRLLQEAREVGAIVVSGVE
RcSDH4        MILANASAVGMEPNTH-QSPVSKEALKSYELVFDAVYTPRNTRLLQEAKEVGAVVVSGVE
MeSDH4        MILANASAVGMEPNSD-QTPVSKEALKAYELVFDAVYTPRNTRLLQEAKEVGAIVVSGLE
SiSDH2        AILANATPLGMHPNTD-RIHVAEETLRDYKLVFDAVYTPRKTRLLKEAEAAGAIVVSGVE
PvSDH1        AILANATPVGMHPNTD-RIPVAEGTLEDYWLVFDAVYTPRRTRLLKEADGAGATTVGGVE
GmSDH2        AILANATPVGMHPNTD-RIPVAEATLEDYRLVFDAVYTPRRTRLLNEADAAGAITVAGVE
NtSDH2        AILANATPIGMHPNKD-RIPVSEASLKDYVVVFDAVYTPRKTTLLKDAEAAGAITVSGVE
StSDH3        AILANATPIGMHPNKD-RIPLPEGSLKDYVVVFDAVYTPRKTTLLEDAEAAGAVIVSGVE
SlSDH2        AILANATPIGMHPNKD-RIPVPEGSLKDYVVVFDAVYTPRRTTLLEDAEAAGALIVSGVE
EgSDH4        SILANATPIGMHPNKD-RIPVSEVPYLK*--------------------------------
TcSDH1        AILANATPLGMHPKTDQRIPVAESTLGDYELVFDAVYTPRKTRLLKDAEAAGAIIVGGVE
GrSDH2        AILANATPLGMHPYTDQRIPVDQETLGDYELVFDAVYTPRKTRLLKEAEAAGAITVSGVE
CasSDH1       AILANATPLGMHPNTD-RIPVAKGTLGDYTVVFDAVYTPRKTTLLKEAEAAGAIIVSGVE
VvSDH7        AILANATPVGMHPNTD-RIPVAEETLSDYQLVFDSVYTPRKTRLLKEAEAAGAIIVSGVE
CsSDH1        AILANATPLGMHPNTD-RVPVSEETLRDYQLVFDAVYTPRKTRLLKDAEAAGAIIVSGVE
CcSDH2        AILANATPLGMHPNTD-RVPVSEETLRDYQLVFDAVYTPRKTRLLKDAEAAGAIIVSGVE
SpSDH1        AILANATPIGMHPSTD-RIPAAEETLGNYQLVFDAVYTPRKTRLLKDAAAAGAITVSGVE
Poptr3        AILANATPIGMHPSTD-RIPAAEETLGNYQLVFDAVYTPRKTRLLKDADAAGAITVSGVE
PeSDH2        AILANATPIGMHPSTD-RIPAAEETLGNYQLVFDAVYTPRKTRLLKDADAAGAITVSGVE
Poptr2        AILANATPIGMHPSTD-RIPVAEATLGNYQLVFDAVYTPRKTRLLEDADAAGAITVSGVE
SpSDH5        AILANATPIGMHPSTD-RIPVAETTLGSYQLVFDAVYTPRKTRLLKDADAAGAITVSGVE
RcSDH2        AILANATPIGMHPNTE-RIPVAEATLGIYQLVFDAVYTPRKTRLLKEAEAAGAIIVSGVE
JcSDH1        AILANATPLGMHPNTE-RIPVAEETLGIYQLVFDAVYTPRKTRLLKEAEAAGAIIVSGVE
SpSDH6        MVLANTTSVGMTPHTD-ATPLAKKKE--------------------KKKKTGAIVVHGTE
TcSDH4        MVLANTTSVVMKPNID-ATPIPKHALKHCCLVFDAIYTPGFYKKR---KTLGRL------
CasSDH2       MILANTTSVGMKPNID-LTPISKEALKHYDLVFDAIYTPKDTRLLREARECGKIIVYGTE
VvSDH8        MILANTTSVGMKPKIN-DTPIPKHALKHYSLVFDAIYTPKDTRLLREAKESGKIIVYGTE
FvSDH1        MILANTTSVGMKPNVD-DTPISKQALKHYCLVFDAIYTPKETRLLREAKETGAAVVYGTE
EgSDH3        MVLANTTSVGMKPKID-ETPLAKHALKNYCLVFDAIYTPKDTRLLREARETGAVIVYGTE
MeSDH3        MVLANTTSVGMKPNFD-ATPLAQHALKHYCLVFDAIYTPKDTRLLREAKESGALVVYGTE
RcSDH1        MVLANTTSVGMKPNID-ATPLAKHALKHYCVVFDAIYTPKDTRLLREAKESGAVIVYGTE
JcSDH3        MVLANTTSVGMKPNYD-ATPLPKHVLKHYCLAFDAIYTPKDTRLLREAKESGAVIVYGTE
SbSDH2        MILANTTAIGMHPNVN-ETPLSKQALKSYAVVFDAVYTPKETRLLREAAECGATVVSGLE
ZmSDH2        MILANTTAIGMHPNVN-ETPLSKQALRSYAVVFDAVYTPKETRLLREAAECGATVVSGLE
SeiSDH2       MILANTTAIGMHPNVN-DTPLSKQALKSYAVVFDAVYTPKETRLLREAAECGATVVSGLE
PavSDH2       MILANTTAIGMHPNVN-ETPLSKQALKSYAVVFDAVYTPKETRLLHEAAECGATVVSGLE
PhSDH2        MILANTTAIGMHPNVN-ETPLSKQALKSYAVVFDAVYTPKETRLLREAAECGATVVSGLE
PavSDH4       MILANTTAIGMHPNVN-ETPLSKQALKSYAVVFDAIYTPKETRLLREAAECGATVVSGLE
OsSDH1        MILANTTAIGMHPNVN-ETPLSKQALKSYAVVFDAVYTPKETRLLREAAECGATVVSGLE
BdSDH2        MVLANTTAIGMHPNVN-DTPLSKQALKSYAVVFDAVYTPKETRLLREAAECGATVVSGLE
TaSDH1        MILANTTAIGMHPNVN-ETPLSKQALKSYAVVFDAVYTPKETRLLKEAAECGATVVSGLE
TaSDH2        MILANTTAIGMHPNVN-ETPLSKQALKSYAVVFDAVYTPKETRLLREAAECGATVVSGLE
ZmSDH1        MILANATSVGMHPNVD-ETPLSKHALKSYAVVFDAVFVPRETRLLREAAVCGATVIDGLE
SbSDH1        MILANATSVGMHPNVD-GIPLSKQALKSYAVVFDAVFVPRETRLLREAAACGATVIDGLE
SeiSDH1       MILANATSVGMYPNVN-ETPLSKEALRNYSVVFDAVYIPKETRLLREAAECGTTVVNGLE
PavSDH3       MILANATSVGMYPNVN-ESPLSKKALRNYSVVFDVVFIPKETRLL*--------------
PhSDH1        MILANATSVGMFPNVN-DTPLSKKALRNYSIVFDAVYIPKETRLLREAAECGATVVDGLD
PavSDH1       MILANATSVGMYPNVN-ETPLSKKALRNYSIVFDAVYIPKETRLLREAAECGATVVDGLD
SeiSDH3       MILANATSLGMYPNID-GTPIPKKSLSFYNVVFDAVYAPKVTRLLREAEECGVKVVSGVE
PhSDH3        MILANATSLGMQPNID-GTPIPKKALSFYDVVFDAVYAPKVTRLLREAEECGVKVVSGVE
SbSDH3        MILANATSLGMYPNVD-GTPIPKKALGFYDVVFDAVYAPKVTRLLREAEECGVKVVSGVE
ZmSDH3        MILANATSLGMYPNVD-GTPIPKKALGFYDVVFDAVYAPKVTRLLREAEECGVKVVSGVE
TaSDH3        TILANATSLGMYPNVD-GTPVPKKALRFYDVVFDAVYAPKVTRLLREAKEHGVKVVSGVE
OsSDH2        MILANATSLGMYPNVD-GTPIPKKALSFYDVVFDAVYAPKVTRLLREAEECGIKVVSGVE
BdSDH1        MILANATSLGMYPNVD-GTPIPKKALSSYDVVFDAVYAPKVTRLLREAEEQGIKVVSGVE
MaSDH3        MILANATSVGMQPNVG-ETPLAKRALGHYALVFDAVYAPKVTRLLREAEECGVPTVGGFE
MaSDH4        MILANATSVGMQPNVG-ETPLAKRALGHYALVFDAVYAPKVTRLLREAEECGVPTVGGFE
ElgSDH2       MILANTTSVGMRPNVN-DTPVSKQALRSYAVVFDAVYTPKVIRLLREAKESGAAVVSGVE
ElgSDH1       MILANTTSIGMHPNVI-DTPLSKQALRSYAVVFDAVYTPKVTRLLREAEESGAAVVSGLE
PdSDH1        MILANTTSIGMHPNVN-DTPLSKQALRFYAVVFDAVYTPKVTRLLREAKESGAAVVSGVE
PdSDH2        MILANTTSIGMHPNVK-DTPLSKQALKYYAVVFDAVYTPKVTRLLREAKEAGAAVVSGEE
MaSDH1        MILANTTSIGMQPNVD-ETPLKKQALGAYAVVFDAVYTPKVTRLLREAEESGATVVSGLE
MaSDH2        MILANTTSIGMQPNVN-GTPIRKQALGSYSVVFDAVYTPRATRLLREAEESGATVVSGLE
PvSDH2        MILANATSVGMKPKNH-ESLVPKEALKHYSLVFDAVYTPKLTRLLQEAREVGAAIVYGTE
GmSDH1        MILANTTSVGMKPKID-LTPIPKEALKHYSLVFDAIYTPKLTRLLREAQETGAAIVYGTE
FvSDH2        MVLANTTSVGMKPKTD-QTPIPKEALKNYCLVFDAIYTPKWTRLLTEAQESGAAVVFGTE
```

```
PpSDH1       MVLANTTSVGMKPRTD-QTPISKKALKHYSLVFDAIYTPKWTRLLQEAQESGAAVVSGTE
PmSDH3       MVLANTTSVGMKPRTD-QTPVSKKALKHYSLVFDAIYTPKWTRLLQEAQESGAAVVFGTE
MdSDH1       MVLANTTPVGMKPRID-QTPLSKNALKNYSLVFDAIHTPKWTRLLQEAQESGAAVVSGTE
PbSDH1       MVLANTTSVGMKPRID-QTPLSKNALKNYSLVFDAILTPKWTRLLQEAQESGAAVVSGTE
MdSDH4       MVLANATSVGMEPRID-QTPLSKNTLKNYSLVFDAIYTPKWTRLLQEAQESGASVAFGTE
PbSDH4       MVLANATSVGMEPRID-QTPLSKNALKNYALVFDAIYTPTWTRLLQEAQECGAAVAFGTE
TcSDH3       MILVNTTSVGMKPKID-ETPVSKEALKHYSLVFDAVCTPKLTRLLREAQESGATIVYGAE
GrSDH3       MILVNTTPVGMEPRIE-ETPVSKETLKHYSLVFDAVYTPKLTRLLREAKQSGATIVYGTE
CasSDH3      MILANTTSVGMKPKTD-ATPISKKALNRYSLVFDAIYTPKWTRLLQEAQDSGAKVVLGTE
DkSDH        MILANATSVGMKPNTD-ATPISKEALSRYSLVFDAIYTPKWTRLLREAKETGAKVVFGTE
EgSDH2       MVLANTTSVGMKPNVD-LTPLPKNALSRYCLVFDAIYTPKLTRLLREAQEVGAIPVYGTE
VvSDH9       MILANTTSVGMKPNID-NTPLSKKALSRYSLVFDAIYTPKLTRLLREAQESGAIIVYGTE
SpSDH2       TILANATSVGMKPRIE-DTLLPQEALKHYALVFDAVYTPKLTTLLREAQEAGSTIVYGTE
Poptr5       MILANTTSVGMKPRIE-DTPLAKEALKHYALVFDAIYTPKLTTLLREAQEAGSTIVYGTE
PeSDH3       MILANTTSVGMKPRIE-DTPLAKEALKHYALVFDAIYTPKLTTLLREAQEAGSTIVYGTE
MeSDH1       MILANTTSVGMKPRIH-DTPLPKEALKHYSLVFDAIYTPKMTRLLTEAQESGATIVYGTE
JcSDH4       MILANTTSVGMKPRID-DTPLPKDALKQYSLVFDAIYTPKLTRLLREAQESGATIVYGTE
CpSDH        ------------------------------------------------------------
AtSDH        MVLANTTSMGMQPNVE-ETPISKDALKHYALVFDAVYTPRITRLLREAEESGAITVSGSE
AlSDH        MVLANTTSMGMQPNVD-ETPISKHALKHYALVFDAVYTPRITRLLREAEESGAITVSGSE
PpSDH2       MILANTTSIGMQPKVD-DTPISKHALRSYSLVFDAIYTPRVTRLLREAAECGVTVVSGLE
PmSDH2       MILANTTSIGMQPKVD-DTPISKHALRSYSLVFDAIYTPRVTRLLREAAECGVTVVSGLE
PmSDH1       MILANTTSIGMQPKVD-DTPISKHALRSYSLVFDAIYTPRVTRLLREAAECGVTVVSGLE
MdSDH2       MILANTTSIGMQPNVG-DTPLSKHALKYYSLVFDAIYTPKMTRLLREAEECGVTIVSGLE
PbSDH2       MILANTTSIGMQPNVG-DTPLSKHALKYYSLVFDAIYTPKMTRLLREAEECGVTIVSGLE
MdSDH5       MILANTTSIGMQPKVD-ETPISKHALGSYSLVFDAVYTPKMTRLLKEAKEAGATVVSGLE
FvSDH3       AILANSTSVGMQPNID-ETPIPKHALRSYSLVFDAVYNPRMTRLLSEAAESGVRVVCGVE
FsSDH        MILANTTSIGMQPKVD-ETPIPKHALRSYSVVFDAVYTPKITRLLRESEESGAIIVTGVE
JrSDH        MILANSTSIGMQPKVD-ETPIPKHALRSYSLVFDAVYTPKMTRLLREAEESGAKIVTGLE
FvSDH4       MILANTTSIGMQPKVD-ETPISKHALRSYTLVFDAVYTPKITRLLREAEESGAIVVSGSE
PpSDH3       MILANTTSIGMQPKVD-ETPISKHALRSYSLVFDAVYTPKMTRLLKDAEASGAIVVSGLE
MdSDH3       MILANTTSIGMQPKVD-ETPISKHALRSYSLVFDAVYTPKMTRLLKEAKESGATVVSGLE
PbSDH3       MILANTTSIGMQPKVD-ETPISKHALRSYSLVFDAVYTPKMTRLLKEAKESGAIVVSGLE
NnSDH1       MILANTTSIGMQPKVD-ETPIPKKALSSYSLVFDAVYTPKITRLLREAEESGVTIVTGLE
NnSDH2       MILANATPIGMQPKVD-ETPIPKKSLRFYSLVFDAVYTPKITRLLREAGESGITVVTGVE
SiSDH3       MILANTTSIGMQPKVE-ETPVPKEALRHYKLVFDAVYTPKITRLLREAEEMGAKIVTGVE
NtSDH1       MILANTTSIGMQPKVD-DTPIFKEALRYYSLVFDAVYTPKITRLLREAHESGVKIVTGVE
StSDH2       MILANTTSIGMQPKVD-ETPISKEALKHYSLVFDAVYTPKITRLLREAQESGAKIVTGVE
SlSDH1       MILANTTSIGMQPKVD-DTPISKEALKHYSLVFDAVYTPKITRLLREAQESGAKIVTGVE
MtSDH        MILANTTSIGMQPKVD-ETPISKHALKFYSLVFDAVYTPKMTRLLKEAEESGATIVTGLE
PvSDH3       MILANTTSIGMQPKVD-DTPISKHALKYYTLVFDAVYTPKITRLLKEAEESGATIVTGLE
GmSDH3       MILANTTSIGMQPKVD-ETPVSKHALKYYSLVFDAVYTPKITRLLKEAEESGATIVTGLE
LuSDH1       MILANTTSIGMEPKVD-ETPVPKRALRHYALVFDAVYTPKITRMLREAEEWGATVVSGLE
LuSDH2       MILANTTSIGMEPKID-ETPIPKHALRHYALVFDAVYTPKITRMLREAEECGATIVSGLE
TcSDH2       MILANTTSIGMQPNID-DTPIPKDALKYYSLVFDAIYTPKITRLLREAEESGATIVSGVE
GrSDH1       MILANTTSIGMQPNID-QTPMPKDALKYYSLVFDAVYTPKITRLLREAKETGATIVSGLE
SpSDH4       MILANTTSIGMQPKVD-ETPVSKNALRSYSLVFDAVYTPKITRLLREAEESGAKIVTGLE
Poptr1       MILANTTSIGMQPKVD-ETPVSKNALRSYSLVFDAVYTPKITRLLREAEESGAKIVTGLE
PeSDH1       MILANTTSIGMQPKVD-ETPVSKNALRSYSLVFDAVYTPKITRLLREAEESGAKIVTGLE
MeSDH2       MILANTTSIGMQPKVD-ETPVAKPALRYYSLVFDAVYTPKVTRLLREAEESGAITVSGLE
RcSDH3       MVLANTTSIGMQPKVE-ETPIPKHALRNYSLVFDAVYTPKITRLLREAEESGAATVSGLE
JcSDH2       MILANTTSIGMQPKVD-ETPVPKTALRFYSLVFDAVYTPKITRLLREAEECGAITVSGLE
VvSDH5       MILANTTSIGMQPKVD-ETPISKHALKYYSLVFDAIYTPKITRLLREAQESGATIVTGLE
EgSDH5       MILANTTSIGMQPKVD-ETPIPKHALKHYSLVFDAVYTPKITRLLKEAEECGATIVSGLE
CsSDH3       MILANTTSIGMQPKVD-ETPIPKHALGHYALVFDAVYTPKITRLLREAEESGATIVSGLE
CcSDH1       MILANTTSIGMQPKVD-ETPIPKHALGHYALVFDAVYTPKITRLLREAEESGATIVSGLE
```

```
StSDH1       MFVRQALGQFKLFTNGLAPVDF-MR---RIVYEQF-------------------------
SlSDH3       MFVRQALGQFKLFTNGLAPVDF-MR---RIVYEQF-------------------------
SiSDH1       MFIRQALGQFKLFTGGLGTQLA-LISR--LLVSNLNH---IVICI--------TFAKCVG
EgSDH1       MFIRQALHQFKLFTSGLAPEEF-MR---KLVLEQ-----------------------F*-
GrSDH4       MFVRQALGQFRLFTGGSAPEDL-MR---KLVMEQ-----------------------F*-
CsSDH2       MFIRQALGQFRLFTGGLAPEDF-MR---KLVLEQ-----------------------F*-
CcSDH3       MFIRQALGQFRLFTGGLAPEDF-MR---KLVLEQ-----------------------F*-
Poptr4       MFIRQALGQFRLFTGGLGIILP-VVSTHQLRLND-----------------------QLT
SpSDH3       MFIRQALGQFRLFTGGLAPEAF-MR---KLVLEQ--------------------------
JcSDH5       MFIRQALGQFRLFTGGLAPEAF-MR---KLVLEQ--------------------------
RcSDH4       MFIRQALGQFRLFTGGLAPEAF-MR---KLVLEQ--------------------------
MeSDH4       MFIRQAIGQFRLFTGGLAPEAF-MR---KLVLEQ--------------------------
SiSDH2       MFLRQAAAQFSLFTGHKAPEEF-MRGIIL---EKF-------------------------
PvSDH1       MFLRQAIGQFNLFTGLEGTIFF-FSFHGTFFLSLNSI-----FSY--------SFAL---
GmSDH2       MFLRQAIGQFNLFTGLEAPEEF-MREIV---LSK--------F-----------------
NtSDH2       MFLRQAIGQFHLFTRTKAPEEF-MRDIV-M--AKF-------------------------
StSDH3       MFLRQAIGQFNLFTGSKAPEEF-MREIV-M--SKF-------------------------
SlSDH2       MFLRQAIGQFNLFTGSKAPEEF-MRDIV-M--SKF-------------------------
EgSDH4       ------------------------------------------------------------
TcSDH1       MFLGQAIGQFNLFTGQEAPKKL-MREII-M--AKF-------------------------
GrSDH2       MFLRQAVSQFNLFTGQQAPEEL-MREII-A--SQF-------*-----------------
CasSDH1      MFLRQAIEQFNLFTGGKAPQEF-MRRTIF---ANF-------------------------
VvSDH7       MFLRQAIGQFNLFTGGEAPEEF-MREII-L--SKF-------*-----------------
CsSDH1       MFLRQAIGQFNLFTGKEAPKEF-MREIV-L--AKF-------*-----------------
CcSDH2       MFLRQAIGQFNLFTGKEAPKEF-MREIV-L--AKF-------*-----------------
SpSDH1       MFLRQAIGQFNLFTGREAPKDF-MRKIV-L--AKF-------*-----------------
Poptr3       MFLRQAIGQFNLFTGREAPKDF-MREIV-L--AKF-------------------------
PeSDH2       MFLRQAIGQFKLFTGREAPKDF-MREIV-L--AKF-------------------------
Poptr2       MFLKQAIGQFSLFTGREAPKDF-MR-EIVL--AK--------F-----------------
SpSDH5       MFLRQAIGQFNLFTGREGISSS-HLFHFFL--AEA-------YCM--------SPPL---
RcSDH2       MFLRQAMGQFSLFTGREAPTEF-MREIV-L--AKF-------------------------
JcSDH1       MFLRQAIGQFNLFTGREAPEDF-MREIV-L--AKF-------------------------
SpSDH6       MLIRQGFEQYKNFTGLPAPEQL-FRQLMEEHV*----------------------------
TcSDH4       ------------------------------------------------------------
CasSDH2      MLIRQGFEQYKNFTGLQAPEEL-FRELMSRHA-----------------------------
VvSDH8       MLIRQGFEQYKNFTGLPAPEEL-FRELMSKHA*----------------------------
FvSDH1       MLIRQGFEQYKNFTGLPAPEAL-FRELMEKHA-----------------------------
EgSDH3       MLIRQGFEQYKNFTGLPAPEEL-FRTLMEKHA*----------------------------
MeSDH3       MLIRQGFEQYKNFTGLPAPEEL-FRQLMDKHA*----------------------------
RcSDH1       MLIRQGFEQYKNFTGLPAPEEL-FRQLMEKHA-----------------------------
JcSDH3       MLIRQGFEQYKNFTGLPAPEEL-FRQLMEKHA-----------------------------
SbSDH2       MFIRQAMGQFEHFTGMPVDDYKFRA--IASSLDNPNTMKPRL*------------------
ZmSDH2       MFIRQAMGQFEHFTGMPVDDYKFRA--IASSLDNPNTMKPRL*------------------
SeiSDH2      MFIRQAMGQFEHFTGMPVDDDKFRA--ITSSLDNPNTMKPRL*------------------
PavSDH2      MFIRQAMGQFEHFTGMPAPDRLMRD--IVLTKIG*--------------------------
PhSDH2       MFIRQAMGQFEHFTGMPVDNENFQA--ITSSLDNPNTMKPRL*------------------
PavSDH4      MFIRQAMGQFEHFTGMPVDNKNFRA--ITSSLDNPNTMKPRL-------------------
OsSDH1       MFIRQAMGQFEHFTGMPVVDCKFRA--LANSLDNPDAMKAAGI--TAHKARL*-------
BdSDH2       MFIRQAMGQFEHFTGTPVVDNEFRA--LANSLENPCTAKRTQG--AVPKPRL*-------
TaSDH1       MFIRQAMGQFEHFTGMPAVDGKFRA--LANSLDNPSAVKGDRS--TSLKPRL*-------
TaSDH2       MFIRQAMGQFEHFTGMPAVDDKFRA--LSNSLDNPSAVKGDRS--TSLKPRL*-------
ZmSDH1       MLIRLVMVQFKLFTGGMTAPERLMREAILTKTH*----------------------------
SbSDH1       MLIRLVMVQFKLFTGGMTAPERLMREAILTKTH*----------------------------
SeiSDH1      MLARLAVVQFELFTGGMPAPQRLIHDAMTKTQ*-----------------------------
PavSDH3      ------------------------------------------------------------
PhSDH1       MLVRLVMVQFGLFTGGMPAPQRLMREAILTNTQ*----------------------------
PavSDH1      MLLRLVMVQFGLFTGGMPAPRRLMRDAILTSTQ*----------------------------
SeiSDH3      MFIRQAMGQFEHFTGGIEGVMFSSALVLNSLHAKCFNISDYLY--FLQLLKA*-------
PhSDH3       MFIRQAMGQFEHFTGGIEAPESLMRE----IAAKYT*-------------------------
SbSDH3       MFIRQAMGQFERFTGGIEAPECLMRE----IAAKYT*-------------------------
ZmSDH3       MFIRQAMGQFEHFTGGIEAPESLMRE----IAAKYT*-------------------------
TaSDH3       MFVRQAMGQFEHFTGGIEAPESLMR----EIAAQYT*-------------------------
OsSDH2       MFVRQAMGQFEHFTGGIEAPESLMR----EIAAKYT*-------------------------
BdSDH1       MFVRQAMGQFEHFTGGIEAPESLMR----EIAAKYT*-------------------------
MaSDH3       MFIRQAMGQFKLFTGL-EAPGKEMRELMMKYV------------------------------
MaSDH4       MFIRQAMGQFKLFTGL-EAPGKEMRELMMKYV------------------------------
ElgSDH2      MFIRQAMRQFELFTGL-AAPEGLMRNITSSL-------------------------------
ElgSDH1      MFIRQAMGQFELFTGL-AAPESLMRNIISRL-------------------------------
PdSDH1       MFIRQAMRQFELFTEL-AAPESLMRNIVSSF-------------------------------
PdSDH2       MFIRQAMGQFELFTGL-AAPESLMRNIISSL-------------------------------
MaSDH1       MFIRQAMGQFELFTNL-AAPENLIHEIVKKYT------------------------------
MaSDH2       MFIRQAMGQFELFTNL-TAPENLMRDTVMRYT------------------------------
PvSDH2       MFINQAFVQFERFTKLPAPKQL-MRDVLARNT------------------------------
GmSDH1       MFINQAFMQFEMFTKLPAPKQH-MRDVLARNT------------------------------
FvSDH2       MFLNQAFVQVEKFSGIPANKQL-IRDTLARNT------------------------------
```

```
PpSDH1       MLLNQAFVQIESFSGVPAPKQL-IRDVLARNT----------------------------
PmSDH3       MLLNQAFVQIENFSGVPAPKQL-IRDVLARNT----------------------------
MdSDH1       MFLNQAFVQVEKFSGSPAPKQL-IRDVLARNA----------------------------
PbSDH1       MFLNQAFVQVEKFSGSPAPKQL-IRDVLARNA----------------------------
MdSDH4       MFLNQAFVQVEKFSGSPAPKHL-IRDVLARNT----------------------------
PbSDH4       MFLNQAFVQVEMFSGSPAPKHL-IRDVLARNT----------------------------
TcSDH3       MFINQAFVQFEMFTGLPAPKRL-IRDVLARNT----------------------------
GrSDH3       MFINQAFIQFEMFTGLPAPKQL-IRDVLAITT*----------------------------
CasSDH3      MFINQAFVQFERFTGMPAPKEL-IRETLAKNT----------------------------
DkSDH        MFLNQAFVQFEKFTGLPAPKDL-IRETLARNT----------------------------
EgSDH2       MFINQAFVQFERFTGYPEQLKSLIRSSSSCSRIRAHLLGQIPFEFWIVLSSVNLFCCNNK
VvSDH9       MFINQAFVQFERFTGLPAPKEL-IREVLVRNT*----------------------------
SpSDH2       MLINQAFVQFERFTGLPAPKQL-IRDVLARNT*----------------------------
Poptr5       MFINQAFVQFERFTGLPAPKQL-IRDVLARNT----------------------------
PeSDH3       MFINQAFVQFERFTGLPAPKQL-IRDVLARNA----------------------------
MeSDH1       MFINQAFVQFERFTGFPAPKQL-IRDVLARNT*----------------------------
JcSDH4       MFINQAFVQFERFTGLPAPKQL-IRDVLARNT----------------------------
CpSDH        ------------------------------------------------------------
AtSDH        MFVRQAYEQFEIFTGLPAPKEL-YWQIMSKY----------------------------
AlSDH        MFVRQAYEQFEIFTGLPAPKEL-YWQIMSKY*----------------------------
PpSDH2       MFIGQAYEQFEKFTGLPAPKEL-FRKVMEN----------------------------H
PmSDH2       MFIGQAYEQFEKFTGLPAPKEL-FRKVMEN----------------------------H
PmSDH1       MFIGQAYEQFEKFTGLPAPKEL-FRKVMEN----------------------------H
MdSDH2       MFXGQAYEQFEKFTGLPGLVHG-LQLVEWLLFRVKV----------------------VTF
PbSDH2       MFIGQAYEQFEKFTGLPAPKKL-FRKVMEN----------------------------H
MdSDH5       MFIGQAYEQFERFTGLPGRTQI-GGHIISTKFXGDSGGE--G-----------G---RNG
FvSDH3       MLIGQAYEQFERFTGLPAPKEL-FRKIMDNCL----------------------------
FsSDH        MFIGQAFEQFERFTGLPAPKEL-FKKIMAKY----------------------------
JrSDH        MFIGQAYEQFERFTGLPAPKEL-FRKVMANN----------------------------
FvSDH4       MFIGQAYEQFERFTGLPAPKEL-FRKVVESTS----------------------------
PpSDH3       MFIGQAYEQFERFTGLPAPKEL-FRKVMDNSS----------------------------
MdSDH3       MFIGQAYEQFERFTGLPVQENQ-HIAYRTSCVVRCPSLI--H-----------A--VFSH
PbSDH3       MFIGQAYEQFERFTGLPAPKEL-FRRVMDGSS----------------------------
NnSDH1       MFIRQAYEQYEKFTGLPAPKQL-FREIMATYA----------------------------
NnSDH2       MFIRQAYVQYENFTGLPAPKEL-FRETLATHA----------------------------
SiSDH3       MFIGQAYEQYERFTGLPAPKQL-FKDIVN-------------------------------
NtSDH1       MFIGQAYEQYERFTGLASSKGT-FQENYGWILRARSLS---LF---------NA--ALLV
StSDH2       MFIGQAYEQYERFTGLPAPKEL-FKNIMSTY----------------------------
SlSDH1       MFIGQAYEQYERFTGLPAPKEL-FKNIMSTY----------------------------
MtSDH        MFIGQAYRQYEHYTGLPGKLTH-QKNSSEKLWKTIDESA--S-------------------
PvSDH3       MFMGQAYGQYENFTGLPAPKEL-FRKIMEKY----------------------------
GmSDH3       MFMGQAYGQYENFTGLPAPKEL-FRKIMENY----------------------------
LuSDH1       MFIGQAYQQFERFTGLQAPKEL-FRKTMSK-----------------------------
LuSDH2       MFIGQAYQQFERFTGLPAPKEL-FRKTMSK-----------------------------
TcSDH2       MFIGQAYEQFERFTELPGAKGA-VSENCVKDLANLSMCT--Q--------------FFTI
GrSDH1       MFIGQAYEQFERFTGLPAPKEQ-FRRTMSKY----------------------------
SpSDH4       MFIGQAYEQFERFTELPAPKEL-FQKIM-------SKY*----------------------
Poptr1       MFIGQAYEQFERFTELPAPKEL-FQKIM-------SKY-----------------------
PeSDH1       MFIGQAYEQFERFTELPAPKEL-FQKIM-------SKY-----------------------
MeSDH2       MFIGQAYEQFERFTGLPAPKEL-FRKIM-------CK*----------------------
RcSDH3       MFIGQAYEQFENFTGLPAPKEL-FRKIM-------SKH----------------------
JcSDH2       MFIGQAYEQFERFTGLPAPKEL-FRKIM-------SKY----------------------
VvSDH5       MFIGQAYEQFERFTGLPAPKEL-FKQFISNLQ*----------------------------
EgSDH5       MFIGQAYGQYERYTGLPAPKEL-FR------KI-MSKY*----------------------
CsSDH3       MFIGQAYEQYERFTGLPGKMNA-PHLYKFFVLL-LYSFN--KF---------HIFTYFLF
CcSDH1       MFIGQAYEQYERFTGLPAPKEL-------FQKI-MAKY*----------------------
```

```
StSDH1      ------------------------------------------------------------
SlSDH3      ------------------------------------------------------------
SiSDH1      NFF---PRSSTRRFHAQDCTRAILMQIYRNSRLDKHAQKKTVW-----------------
EgSDH1      ------------------------------------------------------------
GrSDH4      ------------------------------------------------------------
CsSDH2      ------------------------------------------------------------
CcSDH3      ------------------------------------------------------------
Poptr4      DMF---PQNY----------------------------------------------------
SpSDH3      --F---*-------------------------------------------------------
JcSDH5      --F-----------------------------------------------------------
RcSDH4      --F-----------------------------------------------------------
MeSDH4      --F---*-------------------------------------------------------
SiSDH2      ------------------------------------------------------------
PvSDH1      ------------------------------------------------------------
GmSDH2      ------------------------------------------------------------
NtSDH2      ------------------------------------------------------------
StSDH3      ------------------------------------------------------------
SlSDH2      ------------------------------------------------------------
EgSDH4      ------------------------------------------------------------
TcSDH1      ------------------------------------------------------------
GrSDH2      ------------------------------------------------------------
CasSDH1     ------------------------------------------------------------
VvSDH7      ------------------------------------------------------------
CsSDH1      ------------------------------------------------------------
CcSDH2      ------------------------------------------------------------
SpSDH1      ------------------------------------------------------------
Poptr3      ------------------------------------------------------------
PeSDH2      ------------------------------------------------------------
Poptr2      ------------------------------------------------------------
SpSDH5      --S---PDNTS*--------------------------------------------------
RcSDH2      ------------------------------------------------------------
JcSDH1      ------------------------------------------------------------
SpSDH6      ------------------------------------------------------------
TcSDH4      ------------------------------------------------------------
CasSDH2     ------------------------------------------------------------
VvSDH8      ------------------------------------------------------------
FvSDH1      ------------------------------------------------------------
EgSDH3      ------------------------------------------------------------
MeSDH3      ------------------------------------------------------------
RcSDH1      ------------------------------------------------------------
JcSDH3      ------------------------------------------------------------
SbSDH2      ------------------------------------------------------------
ZmSDH2      ------------------------------------------------------------
SeiSDH2     ------------------------------------------------------------
PavSDH2     ------------------------------------------------------------
PhSDH2      ------------------------------------------------------------
PavSDH4     ------------------------------------------------------------
OsSDH1      ------------------------------------------------------------
BdSDH2      ------------------------------------------------------------
TaSDH1      ------------------------------------------------------------
TaSDH2      ------------------------------------------------------------
ZmSDH1      ------------------------------------------------------------
SbSDH1      ------------------------------------------------------------
SeiSDH1     ------------------------------------------------------------
PavSDH3     ------------------------------------------------------------
PhSDH1      ------------------------------------------------------------
PavSDH1     ------------------------------------------------------------
SeiSDH3     ------------------------------------------------------------
PhSDH3      ------------------------------------------------------------
SbSDH3      ------------------------------------------------------------
ZmSDH3      ------------------------------------------------------------
TaSDH3      ------------------------------------------------------------
OsSDH2      ------------------------------------------------------------
BdSDH1      ------------------------------------------------------------
MaSDH3      ------------------------------------------------------------
MaSDH4      ------------------------------------------------------------
ElgSDH2     ------------------------------------------------------------
ElgSDH1     ------------------------------------------------------------
PdSDH1      ------------------------------------------------------------
PdSDH2      ------------------------------------------------------------
MaSDH1      ------------------------------------------------------------
MaSDH2      ------------------------------------------------------------
PvSDH2      ------------------------------------------------------------
GmSDH1      ------------------------------------------------------------
FvSDH2      ------------------------------------------------------------
```

```
PpSDH1      ------------------------------------------------------------
PmSDH3      ------------------------------------------------------------
MdSDH1      ------------------------------------------------------------
PbSDH1      ------------------------------------------------------------
MdSDH4      ------------------------------------------------------------
PbSDH4      ------------------------------------------------------------
TcSDH3      ------------------------------------------------------------
GrSDH3      ------------------------------------------------------------
CasSDH3     ------------------------------------------------------------
DkSDH       ------------------------------------------------------------
EgSDH2      NFPPLLEQGLVKSS*-------------------------------------------
VvSDH9      ------------------------------------------------------------
SpSDH2      ------------------------------------------------------------
Poptr5      ------------------------------------------------------------
PeSDH3      ------------------------------------------------------------
MeSDH1      ------------------------------------------------------------
JcSDH4      ------------------------------------------------------------
CpSDH       ------------------------------------------------------------
AtSDH       ------------------------------------------------------------
AlSDH       ------------------------------------------------------------
PpSDH2      S-----------------------------------------------------------
PmSDH2      S-----------------------------------------------------------
PmSDH1      S-----------------------------------------------------------
MdSDH2      VF----HLGYTHDKD----------------------YEEDAWRL---LEVT--------
PbSDH2      S-----------------------------------------------------------
MdSDH5      KW----PDGYMHREF----------------------GAATAWGG---RWCN--------
FvSDH3      ------------------------------------------------------------
FsSDH       ------------------------------------------------------------
JrSDH       ------------------------------------------------------------
FvSDH4      ------------------------------------------------------------
PpSDH3      ------------------------------------------------------------
MdSDH3      DE----PDC-IRKNV----------------------NQ-LQVGG---DPTA--------
PbSDH3      ------------------------------------------------------------
NnSDH1      ------------------------------------------------------------
NnSDH2      ------------------------------------------------------------
SiSDH3      ------------------------------------------------------------
NtSDH1      TFPPKS-------LH----------------------SCVIAM---------------VL
StSDH2      ------------------------------------------------------------
SlSDH1      ------------------------------------------------------------
MtSDH       ------------------------------------------------------------
PvSDH3      ------------------------------------------------------------
GmSDH3      ------------------------------------------------------------
LuSDH1      ------------------------------------------------------------
LuSDH2      ------------------------------------------------------------
TcSDH2      HFLRRQCDGSTSHL----------------------------------------------
GrSDH1      ------------------------------------------------------------
SpSDH4      ------------------------------------------------------------
Poptr1      ------------------------------------------------------------
PeSDH1      ------------------------------------------------------------
MeSDH2      ------------------------------------------------------------
RcSDH3      ------------------------------------------------------------
JcSDH2      ------------------------------------------------------------
VvSDH5      ------------------------------------------------------------
EgSDH5      ------------------------------------------------------------
CsSDH3      SFGNFSAEGTISENH----------------------GKVLVWSVWSIHYMLLILFSSVI
CcSDH1      ------------------------------------------------------------
```

```
StSDH1      ------------------------------
SlSDH3      ------------------------------
SiSDH1      ------------------------------
EgSDH1      ------------------------------
GrSDH4      ------------------------------
CsSDH2      ------------------------------
CcSDH3      ------------------------------
Poptr4      ------------------------------
SpSDH3      ------------------------------
JcSDH5      ------------------------------
RcSDH4      ------------------------------
MeSDH4      ------------------------------
SiSDH2      ------------------------------
PvSDH1      ------------------------------
GmSDH2      ------------------------------
NtSDH2      ------------------------------
StSDH3      ------------------------------
SlSDH2      ------------------------------
EgSDH4      ------------------------------
TcSDH1      ------------------------------
GrSDH2      ------------------------------
CasSDH1     ------------------------------
VvSDH7      ------------------------------
CsSDH1      ------------------------------
CcSDH2      ------------------------------
SpSDH1      ------------------------------
Poptr3      ------------------------------
PeSDH2      ------------------------------
Poptr2      ------------------------------
SpSDH5      ------------------------------
RcSDH2      ------------------------------
JcSDH1      ------------------------------
SpSDH6      ------------------------------
TcSDH4      ------------------------------
CasSDH2     ------------------------------
VvSDH8      ------------------------------
FvSDH1      ------------------------------
EgSDH3      ------------------------------
MeSDH3      ------------------------------
RcSDH1      ------------------------------
JcSDH3      ------------------------------
SbSDH2      ------------------------------
ZmSDH2      ------------------------------
SeiSDH2     ------------------------------
PavSDH2     ------------------------------
PhSDH2      ------------------------------
PavSDH4     ------------------------------
OsSDH1      ------------------------------
BdSDH2      ------------------------------
TaSDH1      ------------------------------
TaSDH2      ------------------------------
ZmSDH1      ------------------------------
SbSDH1      ------------------------------
SeiSDH1     ------------------------------
PavSDH3     ------------------------------
PhSDH1      ------------------------------
PavSDH1     ------------------------------
SeiSDH3     ------------------------------
PhSDH3      ------------------------------
SbSDH3      ------------------------------
ZmSDH3      ------------------------------
TaSDH3      ------------------------------
OsSDH2      ------------------------------
BdSDH1      ------------------------------
MaSDH3      ------------------------------
MaSDH4      ------------------------------
ElgSDH2     ------------------------------
ElgSDH1     ------------------------------
PdSDH1      ------------------------------
PdSDH2      ------------------------------
MaSDH1      ------------------------------
MaSDH2      ------------------------------
PvSDH2      ------------------------------
GmSDH1      ------------------------------
FvSDH2      ------------------------------
```

```
PpSDH1      -------------------------------
PmSDH3      -------------------------------
MdSDH1      -------------------------------
PbSDH1      -------------------------------
MdSDH4      -------------------------------
PbSDH4      -------------------------------
TcSDH3      -------------------------------
GrSDH3      -------------------------------
CasSDH3     -------------------------------
DkSDH       -------------------------------
EgSDH2      -------------------------------
VvSDH9      -------------------------------
SpSDH2      -------------------------------
Poptr5      -------------------------------
PeSDH3      -------------------------------
MeSDH1      -------------------------------
JcSDH4      -------------------------------
CpSDH       -------------------------------
AtSDH       -------------------------------
AlSDH       -------------------------------
PpSDH2      -------------------------------
PmSDH2      -------------------------------
PmSDH1      -------------------------------
MdSDH2      QSGASILTVAKSSKGVRSFLHYX--------
PbSDH2      -------------------------------
MdSDH5      -XGETIF-----------FCEF---------
FvSDH3      -------------------------------
FsSDH       -------------------------------
JrSDH       -------------------------------
FvSDH4      -------------------------------
PpSDH3      -------------------------------
MdSDH3      -IS----------------------------
PbSDH3      -------------------------------
NnSDH1      -------------------------------
NnSDH2      -------------------------------
SiSDH3      -------------------------------
NtSDH1      DS--SALPFVLRRN------------------
StSDH2      -------------------------------
SlSDH1      -------------------------------
MtSDH       -------------------------------
PvSDH3      -------------------------------
GmSDH3      -------------------------------
LuSDH1      -------------------------------
LuSDH2      -------------------------------
TcSDH2      -------------------------------
GrSDH1      -------------------------------
SpSDH4      -------------------------------
Poptr1      -------------------------------
PeSDH1      -------------------------------
MeSDH2      -------------------------------
RcSDH3      -------------------------------
JcSDH2      -------------------------------
VvSDH5      -------------------------------
EgSDH5      -------------------------------
CsSDH3      QHEASLFIFFFGQKYKRSTCTSILCMEIKS*
CcSDH1      -------------------------------
```

Supplemental Figure 1. Multiple sequence alignment of shikimate dehydrogenases collected from dicot and monocot genomes with Clustal Omega (1.2.1).
Key motifs for substrate and cofactor binding and catalysis in SDH domain are highlighted in green ifsimilar to AtSDH and in red if different to AtSDH.

Dicots:

Arabidopsis thaliana: AtSDH (AAF08579), *Arabidopsis lyrata*: AlSDH (EFH60832.1), *Camellia sinensis*: CasSDH1 (AIZ93902), CasSDH2 (AJA40947), CasSDH3 (AJA40948), *Carica papaya*: CpSDH (evm.model.supercontig_16.98), *Citrus clementina*: CcSDH1 (Ciclev10004711m), CcSDH2 (Ciclev10000874m), CcSDH3 (Ciclev10020342m), *Citrus sinensis*: CsSDH1 (orange1.1g010050m), CsSDH2 (orange1.1g010101m), CsSDH3 (orange1.1g007151m), *Diospyros kaki* : DkSDH (BAI40147), *Eucalyptus grandis*: EgSDH1 (Eucgr.H01214.1), EgSDH2 (Eucgr.H04428.1), EgSDH3 (Eucgr.H04427.1), EgSDH4 (Eucgr.B01770.2), EgSDH5 (Eucgr.J00263.6), *Fagus sylvatica*: FsSDH (ABA54867.1), *Fragaria vesca*: FvSDH1 (XP_004302480), FvSDH2 (XP_004302479), FvSDH3 (XP_004289250), FvSDH4 (XP_004288087), *Glycine max*: GmSDH1 (XP_006606496), GmSDH2 (XP_003520825), GmSDH3 (XP_006573239), *Gossypium raimondii*: GrSDH1 (Gorai.006G075100), GrSDH2 (Gorai.013G116200), GrSDH3 (Gorai.004G250400),

GrSDH4 (Gorai.007G132000), *Jatropha curcas*: JcSDH1 (XP_012089153.1), JcSDH2 (KDP40034.1), JcSDH3 (KDP23584.1), JcSDH4 (KDP23583.1), JcSDH5 (KDP21552.1), *Juglans regia*: JrSDH(AAW65140), *Linum usitatissumum*: LuSDH1 (Lus10033981), LuSDH2 (Lus10012797), *Malus domestica*: MdSDH1 (MDP0000176374), MdSDH2 (MDP0000170005), MdSDH3 (MDP0000279446), MdSDH4 (MDP0000193646), MdSDH5 (MDP0000306786), *Manihot esculenta*: MeSDH1 (cassava4.1_005332m), MeSDH2 (cassava4.1_030256m), MeSDH3 (cassava4.1_004545m), MeSDH4 (cassava4.1_008122m), *Medicago truncatula*: MtSDH (XP_003608198), *Nelumbo nucifera*: NnSDH1 (XP_010244660.1), NnSDH2 (XP_010240930.1), *Nicotiana tabacum*: NtSDH2 (AAS90324), NtSDH1 (AAS90325), *Phaseolus vulgaris*: PvSDH1 (Phvul.006G096400), PvSDH2 (Phvul.007G183800), PvSDH3 (Phvul.003G035000), *Populus euphratica*: PeSDH1 (XP_011010977.1), PeSDH2 (XP_011048878.1), PeSDH3 (XP_011024056.1), *Populus trichocarpa*: Poptr1 (Potri.010G019000), Poptr2 (Potri.013G029900), Poptr3 (Potri.005G043400), Poptr4 (Potri.014G135500), Poptr5 (Potri.013G029800), *Prunus mume*: PmSDH1 (XP_008244632.1), PmSDH2 (XP_008244628.1), PmSDH3 (XP_008246555.1), *Prunus persica*: PpSDH1 (EMJ08566), PpSDH2 (EMJ08564), PpSDH3 (EMJ06228), *Pyrus x bretschneideri* : PbSDH1 (XP_009341996), PbSDH2 (XP_009333763), PbSDH3 (XP_009333759.1), PbSDH4 (XP_009373929.1), *Ricinus communis*: RcSDH1 (EEF49751), RcSDH2 (EEF49750), RcSDH3 (EEF45470), RcSDH4 (EEF41656), *Salix purpurea*: SpSDH1 (SapurV1A.0511s0170.1), SpSDH2 (SapurV1A.0353s0030.1), SpSDH3 (SapurV1A.0053s0210.1), SpSDH4 (SapurV1A.0656s0090.1), SpSDH5 (SapurV1A.0353s0010.1), SpSDH6 (SapurV1A.0353s0020.1), *Sesamum indicum*: SiSDH1 (XP_011079751.1), SiSDH2 (XP_011078708.1), SiSDH3 (XP_011071496.1), *Solanum lycopersicum*: SlSDH1 (AAC17991), SlSDH2 (XP_010327280), SlSDH3 (XP_004242317), *Solanum tuberosum*: StSDH1 (XP_006352799.1), StSDH2 (XP_006342730.1), StSDH3 (XP_006338901.1), *Theobroma cacao*: TcSDH1 (XP_007030118), TcSDH2 (EOY06376), TcSDH3 (EOY10614), TcSDH4 (XP_007030116), *Vitis vinifera*: VvSDH5, VvSDH7, VvSDH8, VvSDH9.

Monocots:

Brachypodium distachyon: BdSDH1 (Bradi2g13170.3), BdSDH2 (Bradi4g05897.1), *Elaeis guinensis*: ElgSDH1 (XP_010914407.1), ElgSDH2 (XP_010914405.1), *Musa accuminata*: MaSDH1 (XP_009386288.1), MaSDH2 (XP_009408239.1), MaSDH3 (XP_009408237.1), MaSDH4 (XP_009408236.1), *Oryza sativa*: OsSDH1 (LOC_Os12g34874.2),OsSDH2 (LOC_Os01g27750.1), *Panicum hallii*: PhSDH1 (Pahal.0129s0053.1),PhSDH2 (Pahal.0129s0052.1), PhSDH3 (Pahal.0012s0288.1), *Panicum virgatum*: PavSDH1 (Pavir.Cb00904.1), PavSDH2 (Pavir.J24113.1), PavSDH3 (Pavir.J15887.1), PavSDH4 (Pavir.J01954.1), *Phoenix dactylifera*: PdSDH1 (XP_008813743.1), PdSDH2 (XP_008813741.1), *Setaria italica*: SeiSDH1 (Si021763m), SeiSDH2 (Si021701m), SeiSDH3 (Si000995m), *Sorghum bicolor*: SbSDH1 (Sobic.008G116100.1), SbSDH2 (Sobic.008G116000.1), SbSDH3 (Sobic.003G169900.1), *Triticum aestivum*: TaSDH1 (Traes_5AS_908BF5D38.1), TaSDH2 (Traes_5DS_F55FF8108.1),TaSDH3 (Traes_3DL_2BA54361D.2), *Zea mays*: ZmSDH1 (GRMZM5G804881_T01), ZmSDH2 (GRMZM2G014376_T01), ZmSDH3 (GRMZM2G314652_T02).

Supplemental Figure 2.

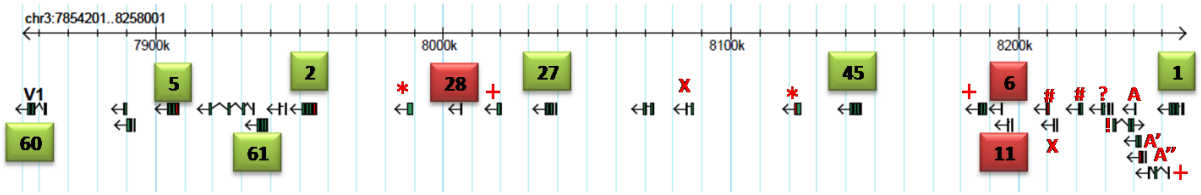
	GAT1	GAT2	VvSCP3	VvSCP4	VvSCP5	VvSCP7	VvSCP8	VvSCP10	VvSCP13	VvSCP14	VvSCP15	VvSCP16	VvSCP18	VvSCP19	VvSCP20	VvSCP21	VvSCP24	VvSCP25	VvSCP26	VvSCP27	VvSCP29	VvSCP31	VvSCP32	VvSCP33
GAT1																								
GAT2	50																							
VvSCP3	24	26																						
VvSCP4	25	27	98																					
VvSCP5	48	53	24	25																				
VvSCP7	19	20	22	22	23																			
VvSCP8	26	29	40	40	28	26																		
VvSCP10	24	24	45	45	25	21	39																	
VvSCP13	26	28	60	61	26	23	39	49																
VvSCP14	27	28	38	39	28	25	41	37	39															
VvSCP15	25	27	38	39	29	26	42	35	40	84														
VvSCP16	28	28	38	39	29	25	43	42	38	47	46													
VvSCP18	18	22	25	25	23	28	21	21	24	25	24	23												
VvSCP19	18	19	20	21	22	95	23	21	21	23	25	23	27											
VvSCP20	26	29	40	41	27	27	85	39	40	42	44	45	21	27										
VvSCP21	25	27	38	38	26	24	47	38	39	47	46	54	22	23	48									
VvSCP24	25	27	59	59	24	26	38	45	67	38	40	38	22	25	39	38								
VvSCP25	35	38	34	32	38	23	29	31	31	32	33	31	23	22	31	33	28							
VvSCP26	24	24	46	48	24	22	35	43	47	38	39	38	22	21	36	36	46	28						
VvSCP27	52	52	23	24	49	21	26	26	29	25	26	28	18	20	24	24	26	34	21					
VvSCP29	23	28	39	40	26	23	52	41	40	44	45	42	23	24	53	45	40	29	37	26				
VvSCP31	46	45	26	27	49	23	30	24	26	28	27	27	22	21	28	26	26	39	23	46	27			
VvSCP32	34	38	29	29	38	23	28	29	33	32	32	28	23	20	28	33	29	64	27	32	29	41		
VvSCP33	27	28	42	42	29	27	50	41	41	41	43	42	23	26	50	44	40	34	37	27	60	28	34	
VvSCP34	20	22	23	23	25	28	20	22	24	25	26	21	85	27	21	23	21	25	22	20	22	23	24	26
VvSCP35	27	26	40	40	27	24	49	41	42	49	50	51	24	24	48	51	40	33	39	25	44	27	34	45
VvSCP36	27	26	37	37	30	24	42	36	38	44	44	51	22	23	44	49	40	31	38	28	42	27	29	42
VvSCP37	26	25	62	63	26	21	38	43	62	35	37	34	22	19	38	35	56	31	43	25	39	27	29	38
VvSCP38	26	29	38	39	25	25	46	38	40	50	50	47	22	24	45	43	39	33	39	26	45	25	32	44
VvSCP39	26	28	37	37	26	25	49	39	38	57	57	50	22	24	48	47	39	33	40	26	48	26	35	47
VvSCP40	26	28	39	39	26	25	79	39	38	38	39	45	22	24	77	45	38	29	34	25	50	28	27	48
VvSCP42	24	23	44	45	24	21	38	75	48	36	36	39	21	20	37	37	45	31	44	22	37	23	29	39
VvSCP44	27	26	37	38	26	25	46	39	39	45	45	52	22	24	46	91	39	34	35	24	43	30	31	42
VvSCP45	52	52	24	24	50	21	26	23	29	26	27	29	18	20	24	25	27	34	21	98	26	45	33	27
VvSCP46	20	21	22	22	23	28	20	22	24	25	24	25	62	27	23	25	25	25	23	18	24	20	26	23
VvSCP47	20	23	23	24	24	31	18	20	24	23	24	23	63	31	20	23	25	23	20	22	24	21	23	21
VvSCP48	20	18	49	49	20	16	28	33	52	26	28	26	16	16	28	29	48	23	33	19	29	21	21	28
VvSCP50	27	26	41	41	29	24	47	41	41	51	50	47	21	24	49	49	41	30	41	23	45	27	28	42
VvSCP51	26	29	39	40	27	25	47	38	37	44	44	45	22	23	49	41	39	29	37	27	48	28	29	44
VvSCP52	46	52	28	28	52	23	27	24	27	26	26	26	23	21	28	26	25	38	25	49	26	50	36	26
VvSCP53	44	49	27	27	44	23	25	24	27	26	26	23	19	21	24	25	26	35	21	46	26	43	36	26
VvSCP54	27	30	40	40	30	23	42	39	40	42	45	47	23	24	42	47	39	29	35	24	41	31	29	40
VvSCP55	26	28	60	59	26	22	38	47	69	39	39	38	21	23	40	38	82	28	46	26	42	26	28	40
VvSCP56	49	54	28	29	52	24	28	24	28	28	28	26	21	23	27	27	29	37	23	50	28	50	36	29
VvSCP57	47	48	26	27	51	23	24	20	25	26	27	25	21	20	24	24	24	35	22	46	25	47	34	25
VvSCP58	24	22	46	47	23	23	37	42	48	38	38	39	24	22	37	36	45	27	82	23	39	23	28	37
VvSCP59	26	25	62	63	26	21	37	44	63	35	38	35	21	21	38	35	57	30	44	24	39	27	28	38
VvSCP60	49	56	26	27	56	22	25	22	28	27	30	27	20	21	28	27	27	39	22	50	24	49	37	25
VvSCP61	49	55	28	27	61	23	29	24	28	29	28	28	21	21	29	27	28	37	23	50	26	49	39	27
VvSCP62	34	38	29	30	38	21	29	29	33	31	31	28	23	20	28	31	29	64	30	34	29	40	94	33
VvSCP63	18	21	17	17	18	24	18	18	18	16	17	16	20	24	18	17	18	16	16	15	19	15	17	19
SlGAC	47	51	26	27	51	22	27	24	28	28	27	26	24	24	27	25	24	36	25	45	28	46	36	27
CtAT1	44	46	19	23	49	21	26	24	26	27	27	26	19	19	26	27	24	41	21	44	27	57	40	26
AtSMT	43	45	26	25	46	24	29	24	28	27	27	26	23	24	30	26	31	41	27	44	28	54	39	29
AtSCT	40	42	27	28	44	23	26	25	29	24	26	25	20	22	25	25	27	35	23	40	22	51	35	27
BnSCT1	39	42	23	25	44	21	25	23	29	25	25	25	19	21	25	25	25	34	24	42	22	50	37	28
BnSCT2	39	42	24	26	44	21	24	23	29	24	25	25	19	21	24	25	24	34	24	42	21	49	37	28
AtSAT	44	46	27	27	47	22	27	25	29	27	27	24	22	22	29	27	28	40	27	43	25	54	38	28
AtSST	44	45	24	25	46	25	28	25	27	25	26	23	25	24	28	23	29	39	24	45	27	55	39	27
AsSCPL1	39	41	25	25	41	21	25	24	26	28	29	23	21	20	24	26	25	33	20	37	25	42	32	23
DkSCPL1	47	61	26	27	49	23	26	26	27	30	30	26	21	21	27	26	24	37	24	45	27	45	37	27
DkSCPL2	58	49	26	26	45	23	27	23	28	28	26	28	17	21	29	26	29	34	27	49	25	43	33	28
AtSCPL17	46	46	25	25	48	24	27	26	29	28	27	25	23	22	27	25	27	40	23	44	25	54	40	25
CsSCPL	64	50	26	25	48	18	26	24	29	28	27	28	17	20	27	28	26	36	21	52	25	45	35	29
SbHNL1	24	23	37	38	25	20	40	35	36	44	44	44	21	18	40	41	34	27	33	22	39	22	26	40
TaCPDWE	26	26	37	37	25	21	48	40	43	55	56	50	24	22	47	47	40	33	40	24	46	24	28	46
HvSCPI	34	37	32	31	37	21	29	29	32	32	34	31	23	20	30	32	27	66	28	34	30	40	61	34
HvCPDWE	27	27	36	36	29	26	45	40	41	69	70	48	27	25	46	45	38	33	39	27	47	28	31	46
HvSCPIII	18	24	20	21	25	30	21	20	23	24	23	23	62	30	19	23	23	25	22	21	24	21	23	27
ScCPY	19	20	22	22	21	24	21	19	26	24	26	24	29	22	21	24	23	23	21	20	22	20	23	20
BRS1	29	28	38	39	27	23	50	40	40	49	50	53	24	22	49	49	38	32	39	26	49	27	30	47
SCPLe	34	39	30	30	39	24	30	30	33	32	34	31	24	23	30	31	29	68	26	34	31	37	62	35
MtSCP1	27	28	40	41	28	24	48	41	40	44	48	47	25	23	51	44	41	29	35	28	65	28	32	56
DgSCPL1	43	46	24	25	47	20	27	23	25	27	28	24	19	20	24	23	23	39	22	44	24	53	37	24
DgSCPL2	45	48	25	27	51	21	29	24	25	30	27	25	22	23	28	23	28	39	28	46	26	52	38	25
DgSCPL3	26	24	38	38	29	23	46	40	39	55	55	46	22	23	46	46	39	33	38	25	45	28	31	46

	VvSCP34	VvSCP35	VvSCP36	VvSCP37	VvSCP38	VvSCP39	VvSCP40	VvSCP42	VvSCP44	VvSCP45	VvSCP46	VvSCP47	VvSCP48	VvSCP50	VvSCP51	VvSCP52	VvSCP53	VvSCP54	VvSCP55	VvSCP56	VvSCP57	VvSCP58	VvSCP59	VvSCP60	VvSCP61
VvSCP35	21																								
VvSCP36	24	47																							
VvSCP37	23	40	35																						
VvSCP38	23	51	44	39																					
VvSCP39	23	54	47	35	63																				
VvSCP40	23	44	41	36	43	44																			
VvSCP42	21	39	36	44	39	40	37																		
VvSCP44	22	50	49	35	43	46	45	36																	
VvSCP45	20	24	26	25	27	26	25	22	24																
VvSCP46	62	24	24	22	25	22	24	21	25	18															
VvSCP47	62	21	23	21	20	21	21	20	22	22	81														
VvSCP48	16	31	28	85	31	27	27	36	30	19	15	15													
VvSCP50	21	63	49	40	50	53	46	39	49	24	23	24	31												
VvSCP51	22	47	43	35	47	48	47	36	40	28	24	24	25	47											
VvSCP52	21	29	26	24	26	25	28	24	26	49	21	23	19	26	25										
VvSCP53	21	27	27	25	24	24	24	25	25	46	22	21	19	24	25	56									
VvSCP54	23	46	48	36	42	46	42	36	45	24	22	22	26	44	41	24	25								
VvSCP55	22	41	38	60	40	40	38	47	38	26	24	24	51	41	38	28	26	40							
VvSCP56	20	30	30	28	28	30	26	22	29	49	22	24	23	25	24	70	57	29	30						
VvSCP57	21	29	25	23	25	25	25	19	24	46	22	21	18	25	23	70	53	28	26	67					
VvSCP58	22	39	38	44	40	39	35	42	35	23	25	22	35	40	37	24	24	37	45	24	23				
VvSCP59	21	40	36	97	40	35	36	44	36	24	21	20	84	40	35	23	24	37	61	28	22	45			
VvSCP60	21	30	28	28	27	26	27	23	28	50	20	21	20	27	31	50	46	28	28	49	47	24	28		
VvSCP61	20	24	29	26	30	26	27	25	28	49	19	22	20	26	29	52	45	28	28	51	51	24	26	57	
VvSCP62	23	32	30	28	32	33	28	29	31	34	26	24	20	28	29	36	36	31	28	36	34	27	28	37	39
VvSCP63	21	18	19	19	17	18	17	20	15	16	18	21	11	17	18	15	8	20	18	16	15	16	20	16	15
SlGAC	22	26	29	25	25	28	26	23	24	45	22	23	18	27	27	47	44	27	24	47	47	24	24	50	51
CtAT1	17	27	26	23	22	23	24	22	27	44	20	22	19	25	25	48	43	25	25	47	46	21	23	48	47
AtSMT	19	26	27	27	28	26	28	25	26	45	21	24	20	25	29	48	43	30	31	47	44	27	27	48	48
AtSCT	20	27	23	24	25	25	24	24	26	41	20	24	18	23	25	44	41	23	27	45	41	24	24	45	43
BnSCT1	21	25	23	23	22	24	25	24	25	42	19	20	18	22	24	45	41	26	24	45	44	25	24	44	45
BnSCT2	21	24	23	23	22	23	24	24	25	42	19	20	17	22	24	46	41	26	24	45	45	24	24	44	45
AtSAT	22	27	27	27	25	24	27	24	29	44	22	24	17	26	28	48	45	27	28	49	45	27	27	48	48
AtSST	23	26	24	24	27	25	28	24	23	45	21	25	17	25	27	49	43	26	26	48	46	24	24	48	48
AsSCPL1	21	27	26	22	23	26	24	23	27	37	23	23	16	27	24	43	38	24	25	40	39	19	22	42	41
DkSCPL1	21	27	26	25	26	26	26	24	26	45	21	20	18	26	27	50	47	28	25	48	48	24	25	53	51
DkSCPL2	19	28	30	28	26	29	28	24	25	49	19	19	24	26	26	46	40	27	28	46	42	25	28	45	45
AtSCPL17	24	26	27	27	30	25	27	24	25	45	22	24	20	27	27	51	46	26	26	48	47	24	26	48	48
CsSCPL	19	27	28	26	24	26	26	22	25	53	20	17	22	28	29	47	45	27	25	46	45	23	26	48	48
SbHNL1	19	46	42	33	43	52	39	35	40	22	22	22	23	43	39	22	22	38	34	23	22	34	34	26	24
TaCPDWII	24	55	47	39	58	63	43	39	47	24	25	25	30	53	47	23	24	47	42	27	24	39	40	28	27
HvSCPI	22	35	31	28	33	33	28	30	32	34	26	24	21	32	32	38	35	30	28	37	34	27	29	38	37
HvCPDWII	28	49	43	37	50	55	42	37	44	27	25	26	28	51	46	29	25	44	40	29	26	41	36	29	27
HvSCPIII	63	25	26	22	24	21	21	21	26	21	66	64	17	24	22	26	21	20	22	23	24	24	21	21	23
ScCPY	27	21	22	21	24	23	20	19	22	19	30	29	16	21	23	19	20	20	22	18	17	22	20	18	21
BRS1	24	71	48	38	51	56	45	40	49	27	23	25	31	61	47	28	24	45	40	29	24	39	38	26	28
SCPLe	24	32	31	29	32	35	27	29	31	34	27	25	21	30	32	36	35	32	30	37	36	29	28	40	39
MtSCP1	25	46	43	38	49	48	46	39	43	28	25	26	28	45	47	25	24	42	42	25	24	39	39	24	26
DgSCPL1	21	25	26	24	27	25	22	21	26	44	20	20	21	23	27	45	41	25	23	47	42	24	23	50	50
DgSCPL2	24	27	25	26	28	24	26	22	26	46	22	22	20	23	27	45	43	26	26	48	45	27	26	51	51
DgSCPL3	22	53	44	36	57	68	42	40	45	26	21	19	27	54	44	26	23	42	39	28	26	41	36	29	28

	VvSCP62	VvSCP63	SlGAC	CtAT1	AtSMT	AtSCT	BnSCT1	BnSCT2	AtSAT	AtSST	AsSCPL1	DkSCPL1	DkSCPL2	AtSCPL17	CsSCPL	SbHNL1	TaCPDWII	HvSCPI	HvCPDWII	HvSCPIII	ScCPY	BRS1	SCPLe	MtSCP1	DgSCPL1	DgSCPL2
VvSCP63	18																									
SlGAC	36	17																								
CtAT1	39	15	44																							
AtSMT	40	14	45	54																						
AtSCT	35	16	44	53	54																					
BnSCT1	37	12	45	53	54	84																				
BnSCT2	37	12	45	53	54	84	98																			
AtSAT	38	10	44	55	76	57	55	55																		
AtSST	39	15	44	55	77	56	54	54	77																	
AsSCPL1	33	17	39	42	40	39	37	37	41	40																
DkSCPL1	38	18	50	45	42	41	41	41	42	43	40															
DkSCPL2	31	14	45	42	42	40	39	39	44	41	36	45														
AtSCPL17	40	13	45	54	60	53	52	52	60	60	38	45	44													
CsSCPL	35	16	46	45	43	41	42	41	45	44	38	48	64	46												
SbHNL1	28	17	26	23	24	23	22	23	23	23	24	25	24	24	23											
TaCPDWII	30	16	28	23	25	23	24	23	24	22	27	25	30	26	25	53										
HvSCPI	63	19	36	40	38	35	36	36	37	38	35	35	33	40	36	30	35									
HvCPDWII	31	16	30	25	29	26	25	25	27	26	28	31	27	26	28	47	54	33								
HvSCPIII	24	18	24	20	22	22	21	21	24	24	24	22	19	25	20	23	22	25	25							
ScCPY	24	21	20	20	21	18	15	15	22	21	19	21	18	21	18	19	26	20	27	26						
BRS1	30	16	27	27	26	25	25	25	25	26	27	27	29	25	27	45	52	33	50	23	21					
SCPLe	63	16	36	39	39	36	36	36	39	38	36	38	33	40	35	28	33	64	36	24	21	31				
MtSCP1	31	19	27	24	27	22	26	25	28	27	23	26	24	26	23	37	48	30	49	24	24	49	32			
DgSCPL1	37	15	45	50	48	48	50	50	48	48	39	46	43	51	44	22	22	39	25	20	18	27	36	24		
DgSCPL2	38	19	48	51	50	47	47	47	50	49	40	47	43	52	46	22	23	37	27	20	18	26	37	25	67	
DgSCPL3	32	19	28	26	28	24	25	24	25	26	26	26	27	27	25	49	61	33	57	21	23	54	34	47	24	24

Supplemental Figure 2. Amino acid identity (%) between serine carboxypeptidases.

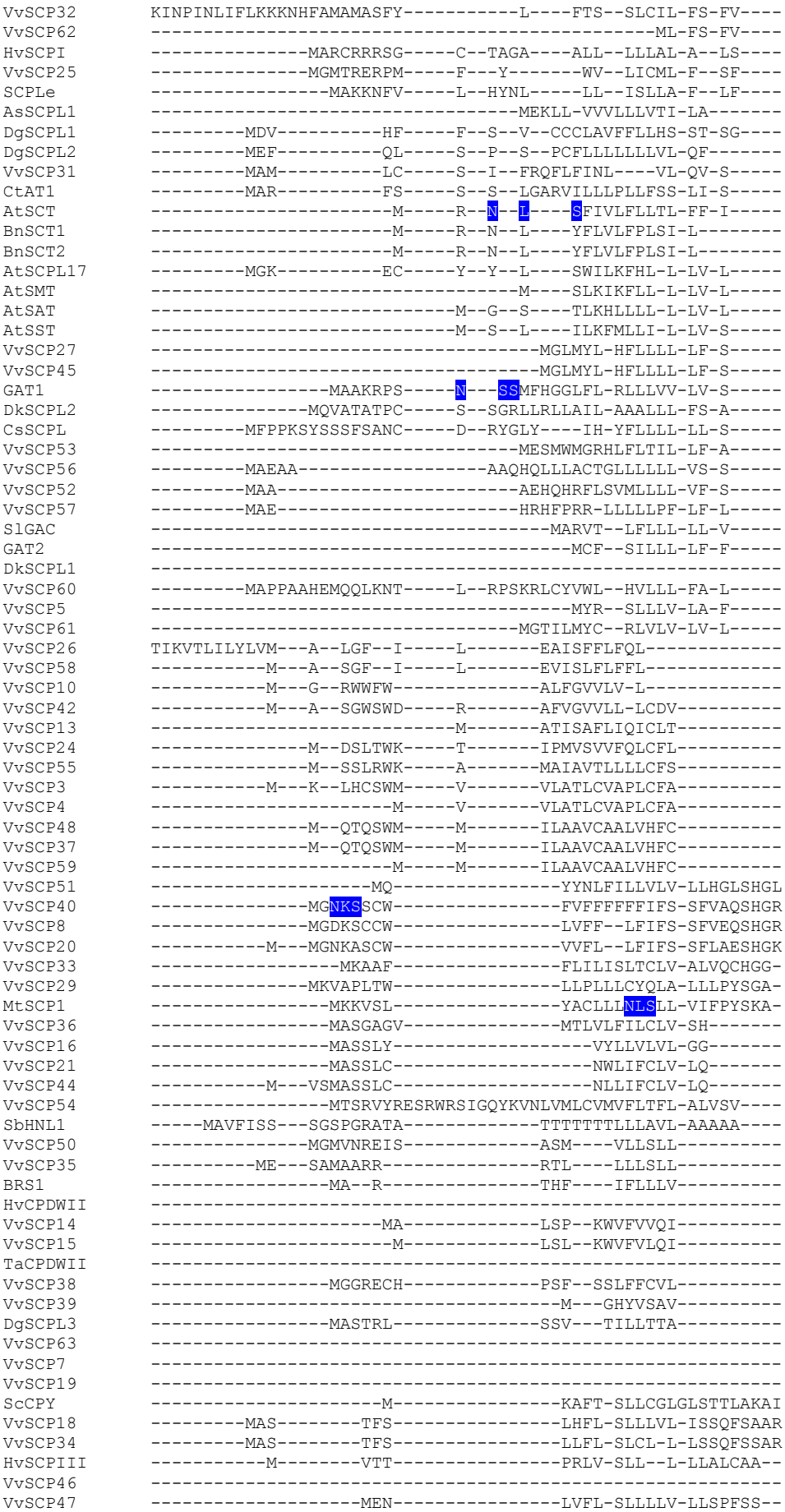
Supplemental Figure 3.

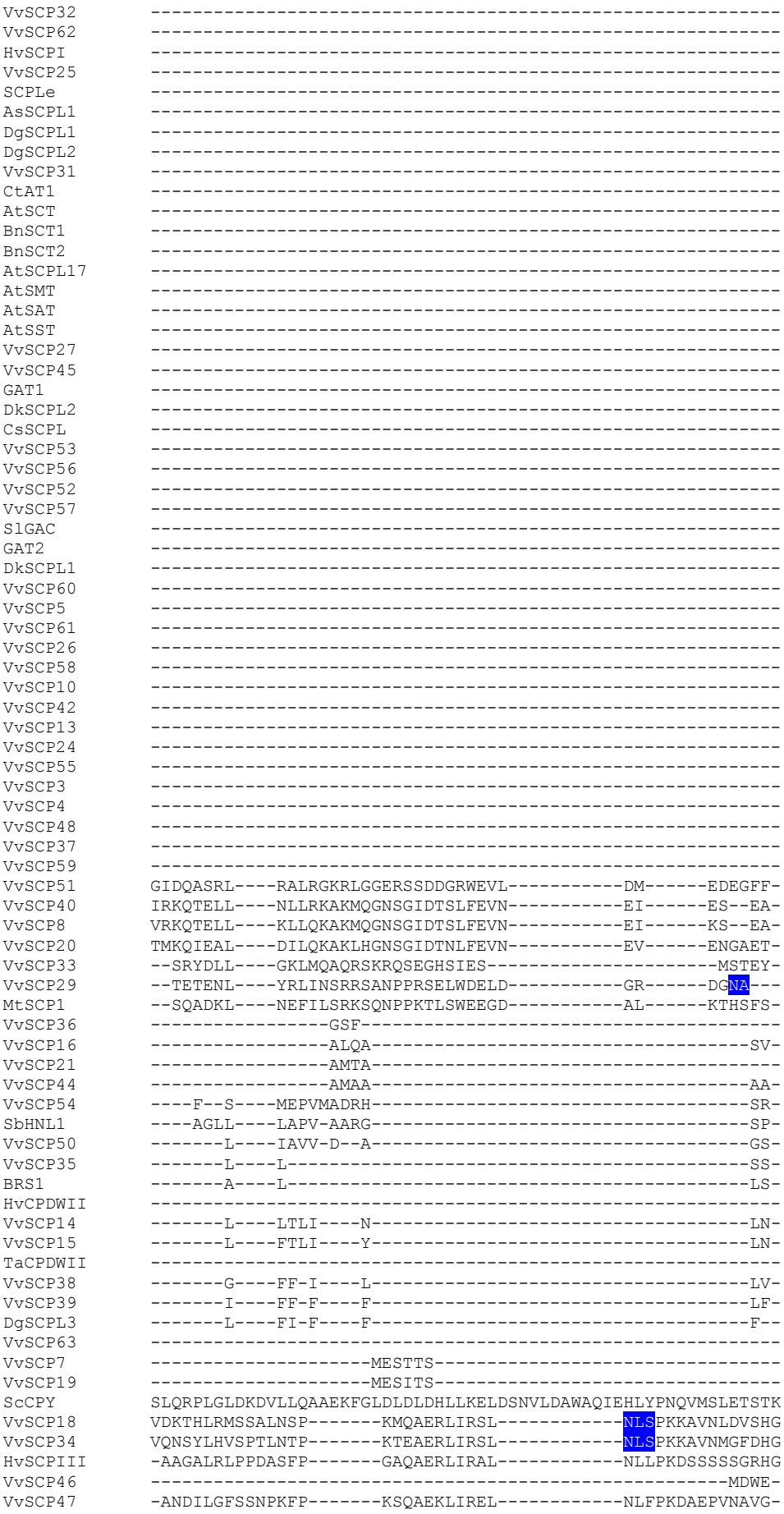


Supplemental Figure 3.Serine carboxypeptidase cluster on chromosome 3 of grapevine genome.
The scale indicates the chromosome and the genomic region shown (http://genomes.cribi.unipd.it/grape/).
Green boxes indicate genes from which a complete proteic sequence has been collected and included in phylogenetic study.
Red boxes indicate genes from which a too short proteic sequence has been collected and not included in phylogenetic study.
The number in the box corresponds to VvSCP or GAT numerotation.
*: Concanavalin A lectin; X: No hit; +: Isoflavone reductase; #: Brassinosteroid insensitive 1-associated receptor kinase 1; ?: Unknown; !: STT3B (staurosporin and temperature sensitive 3-like B); A: Replication factor A 1, rfa1; A': Replication protein A 70 kDa DNA-binding subunit; A'': Replication protein A 70b.

Supplemental Figure 4.

```
VvSCP32       ----------------------------MHIYSVFGLITRPFKIF----------FIFS
VvSCP62       ------------------------------------------------------------
HvSCPI        ------------------------------------------------------------
VvSCP25       ------------------------------------------------------------
SCPLe         ------------------------------------------------------------
AsSCPL1       ------------------------------------------------------------
DgSCPL1       ------------------------------------------------------------
DgSCPL2       ------------------------------------------------------------
VvSCP31       ------------------------------------------------------------
CtAT1         ------------------------------------------------------------
AtSCT         ------------------------------------------------------------
BnSCT1        ------------------------------------------------------------
BnSCT2        ------------------------------------------------------------
AtSCPL17      ------------------------------------------------------------
AtSMT         ------------------------------------------------------------
AtSAT         ------------------------------------------------------------
AtSST         ------------------------------------------------------------
VvSCP27       ------------------------------------------------------------
VvSCP45       ------------------------------------------------------------
GAT1          ------------------------------------------------------------
DkSCPL2       ------------------------------------------------------------
CsSCPL        ------------------------------------------------------------
VvSCP53       ------------------------------------------------------------
VvSCP56       ------------------------------------------------------------
VvSCP52       ------------------------------------------------------------
VvSCP57       ------------------------------------------------------------
SlGAC         ------------------------------------------------------------
GAT2          ------------------------------------------------------------
DkSCPL1       ------------------------------------------------------------
VvSCP60       ------------------------------------------------------------
VvSCP5        ------------------------------------------------------------
VvSCP61       ------------------------------------------------------------
VvSCP26       MIKETLLLINLHASSNSNIKDQYFINKDVVYIYIYKALFSRTRKVQVKSSVFQQEPFPYN
VvSCP58       ------------------------------------------------------------
VvSCP10       ------------------------------------------------------------
VvSCP42       ------------------------------------------------------------
VvSCP13       ------------------------------------------------------------
VvSCP24       ------------------------------------------------------------
VvSCP55       ------------------------------------------------------------
VvSCP3        ------------------------------------------------------------
VvSCP4        ------------------------------------------------------------
VvSCP48       ------------------------------------------------------------
VvSCP37       ------------------------------------------------------------
VvSCP59       ------------------------------------------------------------
VvSCP51       ------------------------------------------------------------
VvSCP40       ------------------------------------------------------------
VvSCP8        ------------------------------------------------------------
VvSCP20       ------------------------------------------------------------
VvSCP33       ------------------------------------------------------------
VvSCP29       ------------------------------------------------------------
MtSCP1        ------------------------------------------------------------
VvSCP36       ------------------------------------------------------------
VvSCP16       ------------------------------------------------------------
VvSCP21       ------------------------------------------------------------
VvSCP44       ------------------------------------------------------------
VvSCP54       ------------------------------------------------------------
SbHNL1        ------------------------------------------------------------
VvSCP50       ------------------------------------------------------------
VvSCP35       ------------------------------------------------------------
BRS1          ------------------------------------------------------------
HvCPDWII      ------------------------------------------------------------
VvSCP14       ------------------------------------------------------------
VvSCP15       ------------------------------------------------------------
TaCPDWII      ------------------------------------------------------------
VvSCP38       ------------------------------------------------------------
VvSCP39       ------------------------------------------------------------
DgSCPL3       ------------------------------------------------------------
VvSCP63       ------------------------------------------------------------
VvSCP7        ------------------------------------------------------------
VvSCP19       ------------------------------------------------------------
ScCPY         ------------------------------------------------------------
VvSCP18       ------------------------------------------------------------
VvSCP34       ------------------------------------------------------------
HvSCPIII      ------------------------------------------------------------
VvSCP46       ------------------------------------------------------------
VvSCP47       ------------------------------------------------------------
```


VvSCP32 KINPINLIFLKKKNHFAMAMASFY----------L----FTS--SLCIL-FS-FV----
VvSCP62 ----------ML-FS-FV----
HvSCPI ----------MARCRRRSG-----C--TAGA----ALL--LLLAL-A--LS----
VvSCP25 ----------MGMTRERPM-----F---Y-------WV--LICML-F--SF----
SCPLe ----------MAKKNFV-----L--HYNL-----LL--ISLLA-F--LF----
AsSCPL1 ----------MEKLL-VVVLLLVTI-LA-------
DgSCPL1 ---------MDV----------HF-----F--S--V--CCCLAVFFLLHS-ST-SG----
DgSCPL2 ---------MEF----------QL-----S--P--S--PCFLLLLLLLVL-QF-------
VvSCP31 ---------MAM----------LC-----S--I--FRQFLFINL----VL-QV-S-----
CtAT1 ---------MAR----------FS-----S--S--LGARVILLLPLLFSS-LI-S-----
AtSCT ----------M-----R--N--L----SFIVLFLLTL-FF-I-----
BnSCT1 ----------M-----R--N--L----YFLVLFPLSI-L---------
BnSCT2 ----------M-----R--N--L----YFLVLFPLSI-L---------
AtSCPL17 ---------MGK----------EC-----Y--Y--L----SWILKFHL-L-LV-L-----
AtSMT ----------M----SLKIKFLL-L-LV-L-----
AtSAT ----------M--G--S----TLKHLLLL-L-LV-L-----
AtSST ----------M--S--L----ILKFMLLI-L-LV-S-----
VvSCP27 ----------MGLMYL-HFLLLL-LF-S-----
VvSCP45 ----------MGLMYL-HFLLLL-LF-S-----
GAT1 ----------MAAKRPS-----N---SSMFHGGLFL-RLLLVV-LV-S-----
DkSCPL2 ----------MQVATATPC-----S--SGRLLRLLAIL-AAALLL-FS-A-----
CsSCPL ---------MFPPKSYSSSFSANC-----D--RYGLY----IH-YFLLLL-LL-S-----
VvSCP53 ----------MESMWMGRHLFLTIL-LF-A-----
VvSCP56 ---------MAEAA----------AAQHQLLLACTGLLLLLL-VS-S-----
VvSCP52 ---------MAA----------AEHQHRFLSVMLLLL-VF-S-----
VvSCP57 ---------MAE----------HRHFPRR-LLLLLPF-LF-L-----
SlGAC ----------MARVT--LFLLL-LL-V-----
GAT2 ----------MCF--SILLL-LF-F-----
DkSCPL1 ----------
VvSCP60 ---------MAPPAAHEMQQLKNT-----L--RPSKRLCYVWL--HVLLL-FA-L-----
VvSCP5 ----------MYR--SLLLV-LA-F-----
VvSCP61 ----------MGTILMYC--RLVLV-LV-L-----
VvSCP26 TIKVTLILYLVM---A--LGF--I-----L-------EAISFFLFQL----------
VvSCP58 ----------M---A--SGF--I-----L-------EVISLFLFFL----------
VvSCP10 ----------M---G--RWWFW----------ALFGVVLV-L----------
VvSCP42 ----------M---A--SGWSWD-----R-------AFVGVVLL-LCDV----------
VvSCP13 ----------M-------ATISAFLIQICLT----------
VvSCP24 ----------M--DSLTWK-----T-------IPMVSVVFQLCFL----------
VvSCP55 ----------M--SSLRWK-----A-------MAIAVTLLLLCFS----------
VvSCP3 ----------M---K--LHCSWM-----V-------VLATLCVAPLCFA----------
VvSCP4 ----------M-----V-------VLATLCVAPLCFA----------
VvSCP48 ----------M--QTQSWM-----M-------ILAAVCAALVHFC----------
VvSCP37 ----------M--QTQSWM-----M-------ILAAVCAALVHFC----------
VvSCP59 ----------M-----M-------ILAAVCAALVHFC----------
VvSCP51 ----------MQ----------YYNLFILLVLV-LLHGLSHGL
VvSCP40 ----------MGNKSSCW----------FVFFFFFFIFS-SFVAQSHGR
VvSCP8 ----------MGDKSCCW----------LVFF--LFIFS-SFVEQSHGR
VvSCP20 ----------M---MGNKASCW----------VVFL--LFIFS-SFLAESHGK
VvSCP33 ----------MKAAF----------FLILISLTCLV-ALVQCHGG-
VvSCP29 ----------MKVAPLTW----------LLPLLLCYQLA-LLLPYSGA-
MtSCP1 ----------MKKVSL----------YACLLLNLSLL-VIFPYSKA-
VvSCP36 ----------MASGAGV----------MTLVLFILCLV-SH-------
VvSCP16 ----------MASSLY----------VYLLVLVL-GG-------
VvSCP21 ----------MASSLC----------NWLIFCLV-LQ-------
VvSCP44 ----------M---VSMASSLC----------NLLIFCLV-LQ-------
VvSCP54 ----------MTSRVYRESRWRSIGQYKVNLVMLCVMVFLTFL-ALVSV----
SbHNL1 -----MAVFISS---SGSPGRATA----------TTTTTTTLLLAVL-AAAAA----
VvSCP50 ----------MGMVNREIS----------ASM----VLLSLL----------
VvSCP35 ----------ME---SAMAARR----------RTL----LLLSLL----------
BRS1 ----------MA--R----------THF----IFLLLV----------
HvCPDWII ----------
VvSCP14 ----------MA----------LSP--KWVFVVQI----------
VvSCP15 ----------M----------LSL--KWVFVLQI----------
TaCPDWII ----------
VvSCP38 ----------MGGRECH----------PSF--SSLFFCVL----------
VvSCP39 ----------M---GHYVSAV----------
DgSCPL3 ----------MASTRL----------SSV---TILLTTA----------
VvSCP63 ----------
VvSCP7 ----------
VvSCP19 ----------
ScCPY ----------M----------KAFT-SLLCGLGLSTTLAKAI
VvSCP18 ---------MAS--------TFS----------LHFL-SLLLVL-ISSQFSAAR
VvSCP34 ---------MAS--------TFS----------LLFL-SLCL-L-LSSQFSSAR
HvSCPIII ----------M--------VTT----------PRLV-SLL--L-LLALCAA--
VvSCP46 ----------
VvSCP47 ----------MEN----------LVFL-SLLLLV-LLSPFSS--

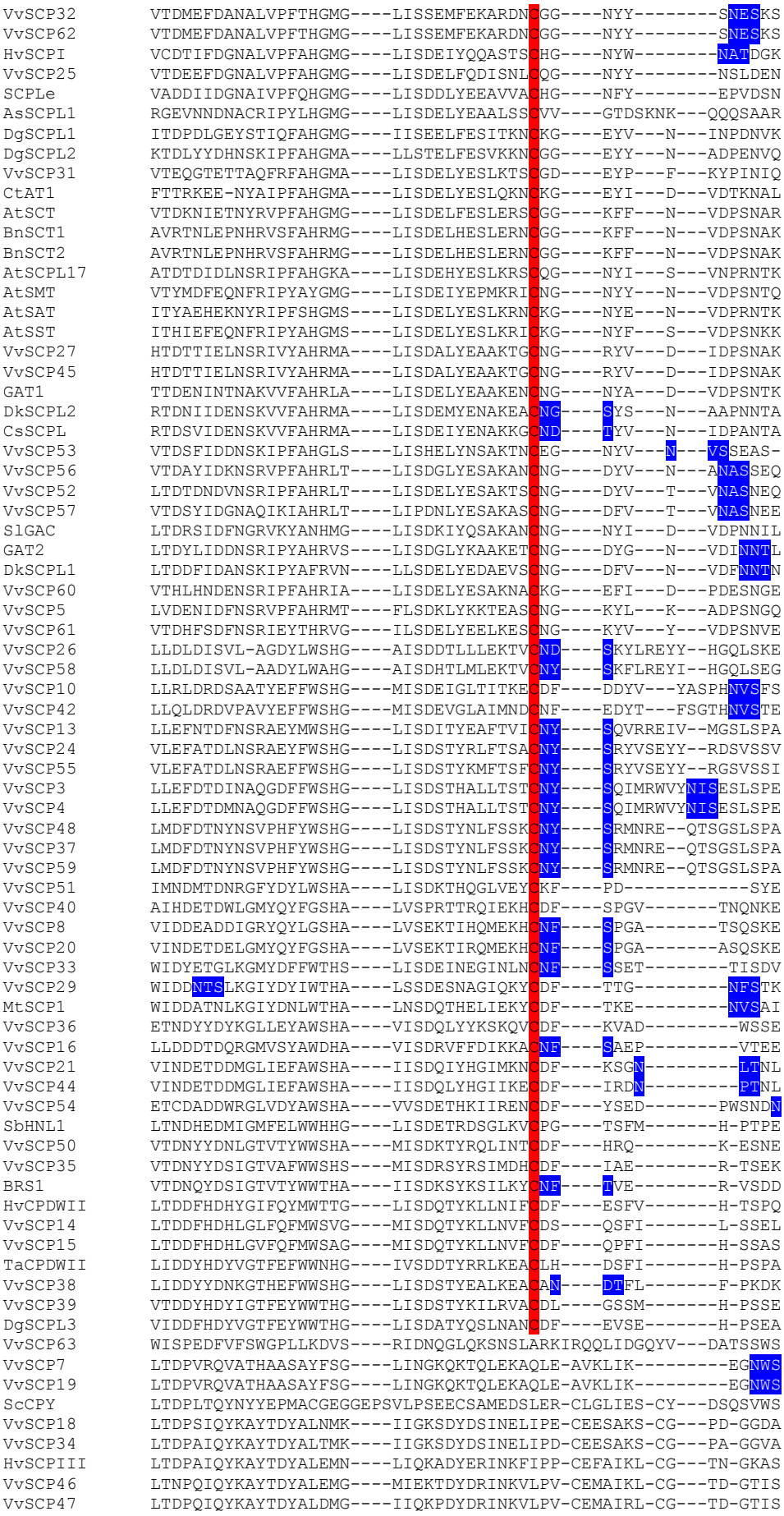

VvSCP32
VvSCP62
HvSCPI
VvSCP25
SCPLe
AsSCPL1
DgSCPL1
DgSCPL2
VvSCP31
CtAT1
AtSCT
BnSCT1
BnSCT2
AtSCPL17
AtSMT
AtSAT
AtSST
VvSCP27
VvSCP45
GAT1
DkSCPL2
CsSCPL
VvSCP53
VvSCP56
VvSCP52
VvSCP57
SlGAC
GAT2
DkSCPL1
VvSCP60
VvSCP5
VvSCP61
VvSCP26
VvSCP58
VvSCP10
VvSCP42
VvSCP13
VvSCP24
VvSCP55
VvSCP3
VvSCP4
VvSCP48
VvSCP37
VvSCP59
VvSCP51 GIDQASRL----RALRGKRLGGERSSDDGRWEVL----------DM------EDEGFF-
VvSCP40 IRKQTELL----NLLRKAKMQGNSGIDTSLFEVN----------EI------ES--EA-
VvSCP8 VRKQTELL----KLLQKAKMQGNSGIDTSLFEVN----------EI------KS--EA-
VvSCP20 TMKQIEAL----DILQKAKLHGNSGIDTNLFEVN----------EV------ENGAET-
VvSCP33 --SRYDLL----GKLMQAQRSKRQSEGHSIES----------MSTEY-
VvSCP29 --TETENL----YRLINSRRSANPPRSELWDELD----------GR------DGNA----
MtSCP1 --SQADKL----NEFILSRKSQNPPKTLSWEEGD----------AL------KTHSFS-
VvSCP36 ----------GSF----------
VvSCP16 ----------ALQA----------SV-
VvSCP21 ----------AMTA----------
VvSCP44 ----------AMAA----------AA-
VvSCP54 ----F--S----MEPVMADRH----------SR-
SbHNL1 ----AGLL----LAPV-AARG----------SP-
VvSCP50 -------L----IAVV-D--A----------GS-
VvSCP35 -------L----L----------SS-
BRS1 -------A----L----------LS-
HvCPDWII ----------
VvSCP14 -------L----LTLI----N----------LN-
VvSCP15 -------L----FTLI----Y----------LN-
TaCPDWII ----------
VvSCP38 -------G----FF-I----L----------LV-
VvSCP39 -------I----FF-F----F----------LF-
DgSCPL3 -------L----FI-F----F----------F--
VvSCP63 ----------
VvSCP7 ----------MESTTS----------
VvSCP19 ----------MESITS----------
ScCPY SLQRPLGLDKDVLLQAAEKFGLDLDLDHLLKELDSNVLDAWAQIEHLYPNQVMSLETSTK
VvSCP18 VDKTHLRMSSALNSP-------KMQAERLIRSL----------NLSPKKAVNLDVSHG
VvSCP34 VQNSYLHVSPTLNTP-------KTEAERLIRSL----------NLSPKKAVNMGFDHG
HvSCPIII -AAGALRLPPDASFP-------GAQAERLIRAL----------NLLPKDSSSSSGRHG
VvSCP46 ----------MDWE-
VvSCP47 -ANDILGFSSNPKFP-------KSQAEKLIREL----------NLFPKDAEPVNAVG-

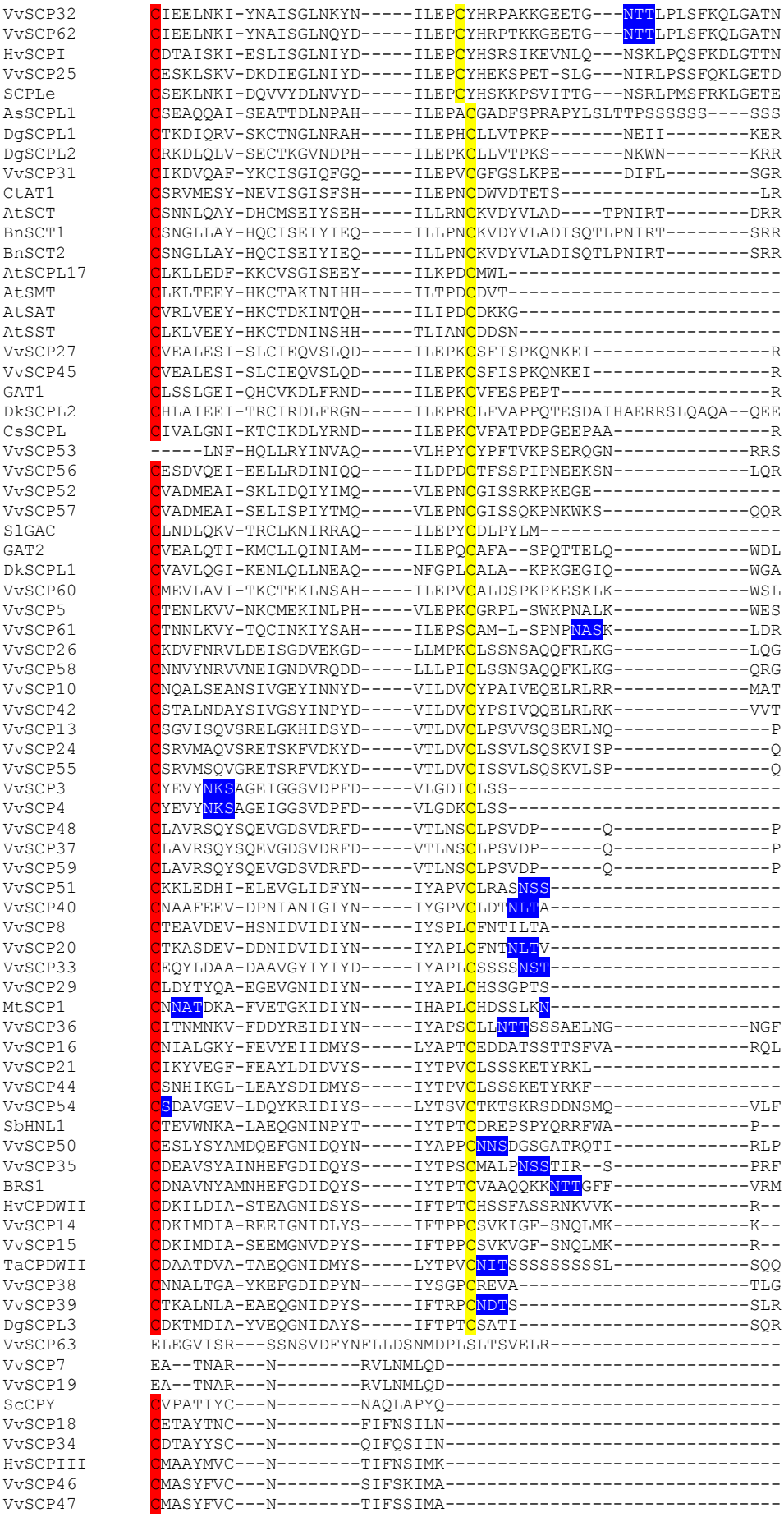
```
VvSCP32       -----------------VFTEAAPK--GSLITHLPGFNGT-------FPSKHYSGYVDIG
VvSCP62       -----------------VFTEAAPK--GSLITHLPGFNGI-------FPSKHYSGYVDIG
HvSCPI        -----------------GGGGAAPQ--GAEVTGLPGFDGA-------LPSKHYAGYVTVD
VvSCP25       -----------------VLTEAAPQ--TALVTKLPGFNGT-------FPSKHYSGYVTID
SCPLe         -----------------LLTEGAPQ--SALVTQLPGFNGT-------FNSKHYAGYVNID
AsSCPL1       -------------------LGAAAE--RTRVTHLKGFDGP-------LPFSLETGYVEVD
DgSCPL1       -----------------SGGGSALA--QSIVRYLPGFDGP-------LPFELETGYTSVV
DgSCPL2       --------------------SSRVHS--QSIVRYLPGFDGP-------LPFHLETGYVSIG
VvSCP31       --------------------SVVAAS--HSPVKFLPGFEGP-------LPFELETGYVGVG
CtAT1         --------------------FQLASC--GTTVDFLPGFDGP-------LPFVLETGYVGVG
AtSCT         --------------------HHLVDA--SLLVKSLPGFEGP-------LPFELETGYVSIG
BnSCT1        ---------------------ILVDA--SLHVKYLPGLEGP-------LPFELETGYVSVG
BnSCT2        ---------------------ILVDA--SLHVKYLPGLEGP-------LPFELETGYVSVG
AtSCPL17      -------------------IQLVDS--GSTIRFLPGFQGP-------LPFELETGYIGVG
AtSMT         --------------------YHHVDS--ASIVKFLPGFEGP-------LPFELETGYIGIG
AtSAT         --------------------IRHVDS--AAIVKSLPGLEGR-------LPFELETGYIGIG
AtSST         --------------------SHHVRS--GSIVKFLPGFKGP-------LPFELETGYIGIG
VvSCP27       --------------------G-KAVS--RNIVKTLPGFSGE-------LPFKLETGYVSVG
VvSCP45       --------------------G-KAVS--RNIVKTLPGFSGE-------LPFKLETGYVSVG
GAT1          --------------------ANAASA--GSIVEYLPGY-GN-------LTFKLETGYVSVG
DkSCPL2       --------------------QPASSA--GQSVKYLPGYDGE-------LPFHLQTGYISVE
CsSCPL        --------------------AQAVLG--GHIVKYLPGYDGE-------LPFKLETGYIRVN
VvSCP53       --------------------SKAVTS--QTIVTSLPGFSGT-------LPFTLETGYVGVG
VvSCP56       --------------------SSVATA--RSIIKTLPGFPGE-------LPFYLETGYVGVG
VvSCP52       --------------------SGIANG--RSVIKTLPGFSGE-------LPFYLETGYVGVG
VvSCP57       --------------------SAIAST--SSIIKTLPGYSGE-------LPFYLETGYVGVG
SlGAC         --------------------YGVVSE--HFIVETLPGFHGK-------LPFTLETGYISVG
GAT2          --------------------TGVATS--QSIVETLPGFPGK-------LPFKLETGYISVG
DkSCPL1       --------------------MAMAAS--PSIVKFLPGFDGE-------LPFKLETGYIGVG
VvSCP60       --------------------FGLSKS--QSIVDTLPGFSGE-------LPFKLETGYVSVG
VvSCP5        --------------------SSIAVS--ESIIKTLPGFEGD-------LPFKLETGYVGVG
VvSCP61       --------------------SSIASS--QSIIKALPGFPGN-------LPFKLETGYVGVD
VvSCP26       --------------------YSKSFA--EL-ITALPGQPAN-------VSFKQYSGYIATD
VvSCP58       --------------------YFKSFA--EL-ITSLPGQPAN-------VSFKQYAGYIVTD
VvSCP10       -----------------SVNGYPEE--DL-VVRLPGQP-E-------VGFRQFAGYVDVD
VvSCP42       --------------CGAANNGWPEE--DL-VVRLPGQP-K-------VGFRQFGGYVDVD
VvSCP13       ---------------VESP-P---SA--DK-IVSLPGQP-Q-------VGFQQFAGYITVD
VvSCP24       --------------LKAH-PSLSHP--DK-IIQLPGQP-Q-------VGFQQFSGYVSLD
VvSCP55       --------------REVE-SSLSLS--DK-ILELPGQP-Q-------VGFQQYSGYVAVD
VvSCP3        --------------MEPV----SES--DK-LGSLPGQP-H-------VSFQQFGGYVTID
VvSCP4        --------------MEPV----SES--DK-LGSLPGQP-H-------VSFQQFGGYVTID
VvSCP48       --------------SSAVESHSAQA--DQ-ISSLPGQP-R-------VT-----------
VvSCP37       --------------SSAVESHSAQA--DQ-ISSLPGQP-R-------VSFQQFSGYITID
VvSCP59       --------------SSAVESHSAQA--DQ-ISSLPGQP-R-------VSFQQFSGYITID
VvSCP51       -----------D-------SGKMED--DLIEGGLPGQPSG-------VLFKQYSGYVTVN
VvSCP40       -----------ESYKILPQKGMKEK--DR-IERLPGQPH--------VGFSQYGGYVTID
VvSCP8        -----------ESYKILPQKGMKEK--DR-IERLPGQPH--------VGFSQYGGYVTID
VvSCP20       -----------ESHKIHPQEGLKEK--DR-IDMLPGQPH--------VGFSQYGGYVTID
VvSCP33       -----------SPVYMGSQDGLKDG--DR-IQALPGQPNG-------LNLDQYSGYVTVD
VvSCP29       -----------SPLYIGPQDGLMQD--DK-IESLPGQPEG-------VNFDQYAGYVTVD
MtSCP1        -----------AAYVAPPQEELRLA--DK-IVTLPGQPYG-------VNFDQYSGYVTVD
VvSCP36       --------------VAGIKAESSQEN--DR-IINLPGQPSS-------PPITQFSGYITVN
VvSCP16       -----------DGHLSHEILARQKA--DR-VKKLPGQPEV--------GFRQYAGYVTVN
VvSCP21       -----------AASVAGAEELQQEK--DR-VKDLPGQPAV--------EFRHYAGYVKLR
VvSCP44       -----------AASVAGAEELQQEK--DR-VKDLPGQPAV--------EFRHYAGYVKLR
VvSCP54       -----------QWTNDKGLNSLGNE--DL-VTDLPGQP-A-------VDFRHYAGYVTVN
SbHNL1        -----------PEHDKQLQLQQQED--DR-IPGLPGQPNG-------VAFGMYGGYVTID
VvSCP50       -----------HGVNGVRGGEEEEA--DR-ITALPGQP-K-------VSFQQYSGYVTVN
VvSCP35       -----------ITSDADAAPKQQSL--DR-ISALPGQP-P-------VTFSQFSGYVTVN
BRS1          -----------TTFPSSSSSREQEK--DR-IKALPGQP-K-------VAFSQYSGYVNVN
HvCPDWII      ------------------------------VPRVPGQAFD-------ASFAHYAGYVTVS
VvSCP14       -----------RATSSSDPLVQQEL--DK-VLQLPGQTFN-------ISFAHYAGYVTVN
VvSCP15       -----------TPASSSDPLVQQRL--DK-VQHLPGQAFN-------ISFAHYSGYVTVN
TaCPDWII      -----------------VEPSGHAA--DR-IARLPGQP-A-------VDFDMYSGYITVD
VvSCP38       -----------SSGATAGNREDQVR--DR-IVKLPGEPPN-------VGFSQYSGYITVD
VvSCP39       -----------VGLCTSSYLEDQER--DR-ITELPGQPKN-------IGFAQYSGYVTVN
DgSCPL3       -------------C---CRAEDEEG--DK-ISELPGQPGN-------VGFSQYSGYVTVD
VvSCP63       ---------MEKWCFLVV---------LFVLMFGPWVNGGMAMAYRSQDGTEEWGYVEVR
VvSCP7        --------------------NLIKLL-LLLLYFLPCFHTKSP-PASTLPLPTKSGYLPVN
VvSCP19       --------------------NLIKLLLLLLLYFLPCFHAKSP-PTSTLPLPTKSGYLPVN
ScCPY         PKFPEAIKTKKDWDFVVKNDAIENYQLRVNKIKDPKILGI------DPNVTQYTGYLDVE
VvSCP18       E--------------------AVAPRMIEKSLEFPFLDGSS--GTSIQDLGHHAGYFRLA
VvSCP34       E----------------------VAPRMVEKSLNFPFLEGSS--GSSVQDLGHRAGYFKLA
HvSCPIII      ARVGE------------GNEDVAPGQLLERRVTLPGLP------EGVADLGHHAGYYRLP
VvSCP46       --------------------GESTKKLVERRLRFPGIDYSDA-S--VKDLSQHAGYYKLR
VvSCP47       --------------------RESPKSLVETRLRFPGIDYSDA-SVSVEDLGHHAGYYKIK
```

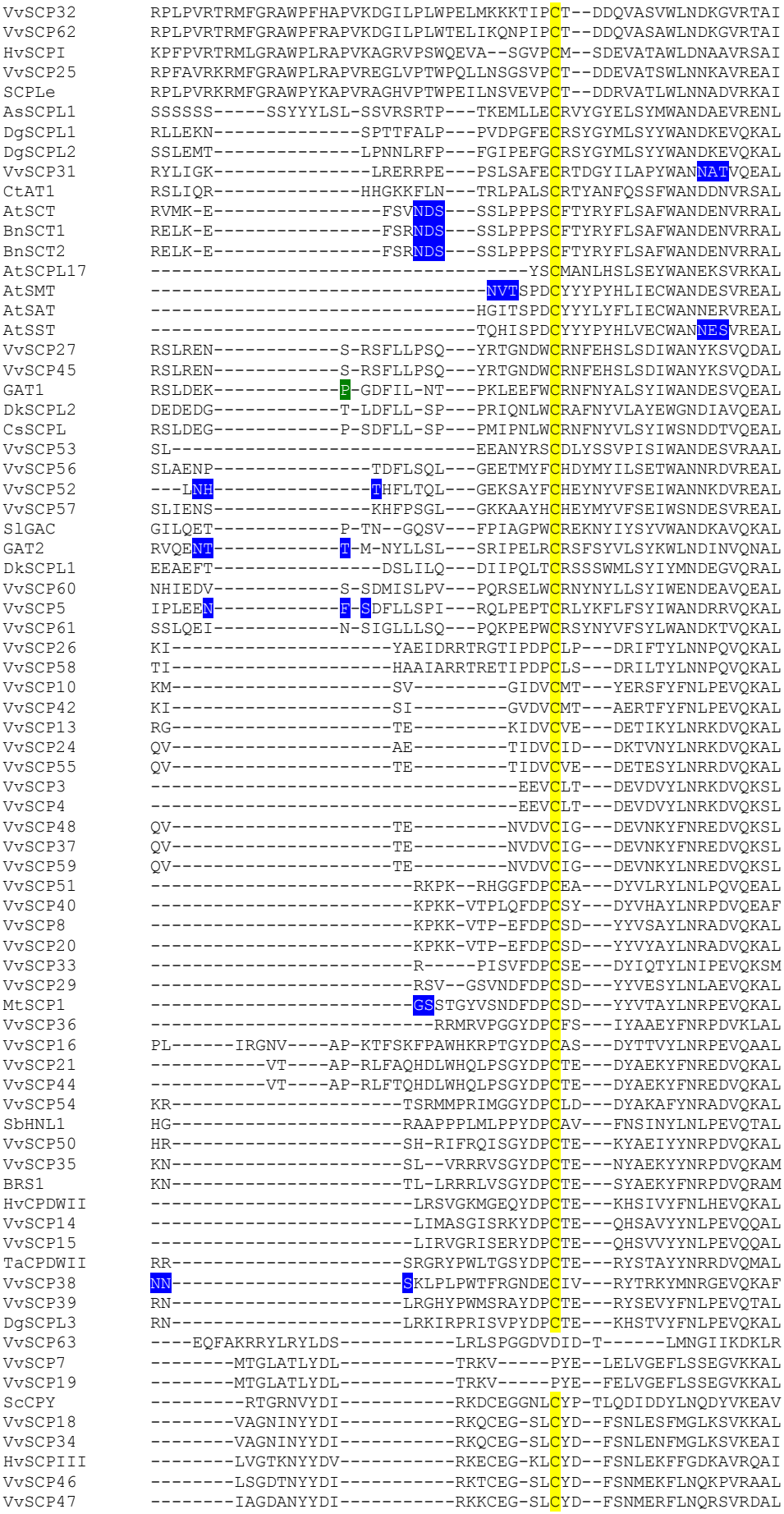
```
VvSCP32        G-EPAKNLFYYFVVS-ER--NP-GKDPLVLWLNGGPGCSSFD-GFVYEHGPFNFEAGKT-
VvSCP62        G-EPAKNLFYYFVVS-ER--NP-AKDPLVLWLNGGPGCSSFD-GFVYEHGPFNFEAGKT-
HvSCPI         E-GHGRNLFYYVVES-ER--DP-GKDPVVLWLNGGPGCSSFD-GFVYEHGPFNFESGGS-
VvSCP25        E-NHGKKLFYYMVVS-EN--NP-SEDPVVLWLNGGPGCSSFD-GFVYEHGPFNFEA-RT-
SCPLe          E-SHGKNLYYYFVES-ER--NP-SKDPVVLWLNGGPGCSSFD-GFVYEHGPFNFDFGKP-
AsSCPL1        E-THGVELFYYFIES-ER--KP-AEDPVILWVSGGPGCSGLN-ALFFEIGPLKLDMASYA
DgSCPL1        DEYGGVEMFYYFISS-ES--KP-SDDPLLFWFTGGPGCSGFS-SIAFENGPVSFKITKY-
DgSCPL2        D-ADGAELFYYFFQS-EN--KP-SDDPILFWFTGGPGCSGLS-AVTFENGPLSFKISKY-
VvSCP31        E-SEEVQLFYYFVKS-EN--NP-TEDPLLLWLTGGPGCSAFS-ALFYEIGPLYFESVPY-
CtAT1          E-GEDVQAYYYFVES-EN--NP-NEDPLMLWLTGGPGCSSFS-GLVLEIGPLIFKREEY-
AtSCT          E-SGDVELFYYFVKS-ER--NP-ENDPLMIWLTGGPGCSSIC-GLLFANGPLAFKGDEY-
BnSCT1         E-SGDVELFYYFVKS-ES--NP-DKDPLMIWLTGGPGCSSIC-GLLFANGPLAFKGDEY-
BnSCT2         E-SGDVELFYYFVKS-ER--NP-DKDPLMIWLTGGPGCSSIC-GFLFANGPLAFKGNEY-
AtSCPL17       E-AEKDQMFYYFIKS-ES--NP-EKDPLLLWLSGGPFCSSFT-ALIYENGPIAFKAEEY-
AtSMT          E-DENVQFFYYFIKS-EN--NP-KEDPLLIWLNGGPGCSCLG-GIIFENGPVGLKFEVF-
AtSAT          E-EEDIQLFYYFIKS-EN--NP-KEDPLLLWLDGGPGCSSLG-GLLFENGPVALKSAVY-
AtSST          E-EENVQFFYYFIKS-DK--NP-QEDPLIIWLNGGPGCSCLS-GLFFENGPLALKNKVY-
VvSCP27        D-I---EFFYYFVES-QG--NP-GADPLILYMNGGPGCSGLN-GFFYQVGPVVFNTTDY-
VvSCP45        D-I---EFFYYFVES-QG--NP-GADPLILYMNGGPGCSGLN-GFFYQVGPVVFNTTDY-
GAT1           D-S---ELFYYFIES-QG--NP-QTDPFFLWLTGGPGCSSFN-GLIYEIGPMEFDIHNY-
DkSCPL2        D-S---ELFYYFIES-EG--NP-LEDPLMLWLTGGPGCSSLY-GIIYEMGPMEFDIHNY-
CsSCPL         E-S---ELFYYFIES-QG--NP-QEDPIFLWLPGGPGCSSFC-ELVYGIGPLEFVIQNY-
VvSCP53        E-SEEVQLFYYFVES-QS--SP-SQDPLMLYIAGGPGCSSLS-SLFYENGPIYLNYQYY-
VvSCP56        E-NESVQLFYYFVKS-QR--NP-VLDPLVLWLTGGPGCSTLS-AFFYESGPVSFNL-TY-
VvSCP52        E-NEEVQLFYYFVKS-QR--NP-VFDPLMLWLSGGPGCSTLT-AFFYENGPLTFNIQEY-
VvSCP57        E-NEEVQLFYMFVKS-QR--NP-VLDPLVMWLTGGPGCSTFS-AFFYGNGPLSFDYKNY-
SlGAC          E-EEKVQLFYFFVQS-ER--DP-RNDPLMIWLTGGPGCSGLS-SFVYEIGPLTFDYANS-
GAT2           D-VDDVQLFYYFIES-ER--NP-RLDPLVLWLTGGPGCSGFS-ALVYEIGPLAFDVEGY-
DkSCPL1        E-IDEVQLFYYFVES-QG--DP-SNDPLILWLTGGPGCSGFS-ALVYEIGPLLFKVQSWK
VvSCP60        E-LNDVELFYYFIES-ER--DP-ARDPLILWLTGGPGCSGFS-GLVYEIGPLRFNYTAF-
VvSCP5         K-SDDIQLFYYFIES-ER--NP-SLDPLMLWLTGGPGCSAFS-GLVYEIGPLIFDYANR-
VvSCP61        D-MDDVQLFYYFVKS-ER--NP-RDDPLLLWLTGGPGCSAFS-GLVYEVGPLSFDYAKS-
VvSCP26        D-QHGRALFYYFVEA-ET--AHPLSRPLTLWLNGGPGCSSLGFGAFMENGPFQPGEN-G-
VvSCP58        A-RHGRALFYYFVEA-KT--ADPLSRPLTLWFNGGPGCSSLGFGAFMENGPFQPGEN-G-
VvSCP10        V-KAGRSLFYYFVEA-ED--DP-DTKALTLWLNGGPGCSSMGGGAFTELGPFFPSGDGR-
VvSCP42        E-KAGRSMFYYFVEA-EE--DP-QNKPLTLWLNGGPGCSSVGGGAFTALGPFFPKGHSR-
VvSCP13        E-KQQRHLFYYFVEA-ET--DP-ASKPLVLWLNGGPGCSSIGAGAFCEHGPFKPSGE---
VvSCP24        D-KKQRALFYYFVEA-ES--DP-ASKPLVLWLNGGPGCSSLGVGAFSENGPFRPNGE---
VvSCP55        E-KQQRALFYYFAEA-ET--DP-AIKPLVLWLNGGPGCSSLGVGAFSENGPFRPSGE---
VvSCP3         E-KQGRALFYYFVEA-VT--DPTASKPLVLWLTGGPGCSSLGGGAFMEHGPFRPRGN---
VvSCP4         E-KQGRALFYYFVEA-VT--DPTASKPLVLWLTGGPGCSSLGGGAFMEHGPFRPRGN---
VvSCP48        ------------------------------------------------------------
VvSCP37        E-KQDRSFFYYFVEA-EN--DTTALKPLVVWFSGGPGCSSVG----AQHGPFRPSGD---
VvSCP59        E-KQDRSFFYYFVEA-EN--DTTTLKPLVVWFSGGPGCSSVGGGAFAQHGPFRPSGD---
VvSCP51        E-LKGRNLFYYFAEA-AE--DP-SSKPLLLWLNGGPGCSSLGVGAMVEIGPFGVKPDGK-
VvSCP40        E-SKGEALYYYFVEA-PT--SK-EYLPLLLWLNGGPGCSSLGAGAMAELGPFRVHSDGK-
VvSCP8         E-SKGEAFYYYFVEA-PT--SR-ESLPLLLWFNGGPGCSSLAYGAMQELGPFRVHSDGK-
VvSCP20        E-SKGKALYYYFAEA-PL--SK-KSLPLLLWLNGGPGCSSLAYGAMQELGPFRVHSEGK-
VvSCP33        P-QAGRALFYYFVES-Q---NS-SSKPLVLWLNGGPGCSSLGSGAMMELGPFRVNGDGN-
VvSCP29        P-KAGRALFYYFVES-PE--DS-STKPLVLWLNGGPGCSSLGYGAMEELGPFRVNPDGK-
MtSCP1         P-EAGRELFYYFVES-PH--NS-YTKPLILWLNGGPGCSSLGYGAFEELGPFRVNSDGK-
VvSCP36        K-AHGRALFYWFFEA-QS--QP-SNRPLLLWLNGGPGCSSIGYGAAVELGPLRVSKNGD-
VvSCP16        E-SHGRALFYWFFEA-TQ--NP-HQKPLLLWLNGGPGCSSIGFGATEELGPFFPRRDG--
VvSCP21        P-QDEKALFYWFFEA-QG--GV-LEKPLVLWLNGGPGCSSIAYGAAQELGPFLVRSNGT-
VvSCP44        P-QDEKALFYWFFEA-QG--GV-LEKPLVLWLNGGPGCSSIAYGAAQELGPFLVRSNGT-
VvSCP54        E-ENGRALFYWFYEA-TT--QP-NEKPLVLWLNGGPGCSSVGYGATQEIGPFIVDTDGH-
SbHNL1         D-NNGRALYYWFQEA-DT-ADP-AAAPLVLWLNGGPGCSSIGLGAMQELGPFRVHTNGE-
VvSCP50        H-VAGRALFYWLNEA-V--HDP-LSKPLVIWLNGGPGCSSVAYGASEEIGPFRINKTAS-
VvSCP35        E-HHGRALFYWLTEA-T--TYP-EKKPLVLWLNGGPGCSSVAYGASEEIGPFRLYRTGS-
BRS1           Q-SHGRALFYWLTES-SS-PSP-HTKPLLLWLNGGPGCSSIAYGASEEIGPFRINKTGS-
HvCPDWII       E-DRGAALFYWFFEA-AH--DP-ASKPLLLWLNGGPGCSSIAFGVGEEVGPFHVNADGK-
VvSCP14        E-YTGRALFYWFIEA-AE--DP-SSKPLVLWLNGGPGCSSIAYGQSEEIGPFHIKEDGK-
VvSCP15        E-NSGRALFYWFIEA-AE--DP-SSKPLVLWLNGGPGCSSIAYGQSEEIGPFHIKEDGK-
TaCPDWII       E-GAGRSLFYLLQEA-PE--DA-QPAPLVLWLNGGPGCSSVAYGASEELGAFRVKPRGA-
VvSCP38        P-RAGRALFYWLIEA-PKSRGP-ASRPLILWLNGGPGCSSVAYGASEEVGPFRVRPDGK-
VvSCP39        K-QAGRALFYWLVQS-PASRGA-ESRPLVLWLNGGPGCSSVAYGAAEEIGPFRIRPDGK-
DgSCPL3        Q-DAGRALFYWLVEA-EA--AA-ATAPLILWLNGGPGCSSIAYGSAEEIGPFRIHPNGR-
VvSCP63        P---KAHMFWWLYRSPYRVESPSKPWPIILWLQGGPGASGVGIGNFQEIGPLGTD-----
VvSCP7         P-TTNSAMFYTFYEA-QNPISPLTQTPLVIWLQGGPGCSS-MIGNFLELGPWRLNRDKHL
VvSCP19        P-TTNSAMFYTFYDA-QNPISPLTQTPLVIWLQGGPGCSS-MIGNFLELGPWRLNCDKHL
ScCPY          D-EDKHFFFWTFESR-NDP----AKDPVILWLNGGPGCSS-LTGLFFELGPSSIGPD--L
VvSCP18        H-SIDARMFYFFFES-RHS----KKDPVVVWLTGGPGCGS-EVALFYENGPFHVRDN--L
VvSCP34        H-TVDARMFYFFFES-RGS----KKDPVVIWLTGGPGCSG-QLALFYENGPFHITDN--L
HvSCPIII       N-THDARMFYFFFES-RGK----KEDPVVIWLTGGPGCSS-ELAVFYENGPFTIANN--M
VvSCP46        H-SLAARMFYLFFES-RDS----RKDPVVIWLTGGPGCGS-ELALFYENGPFTIAAN--M
VvSCP47        H-SSAARMFYLFFES-RDN----RKDPVVIWLTGGPGCSS-ELAVFYENGPFTIAKN--L
```

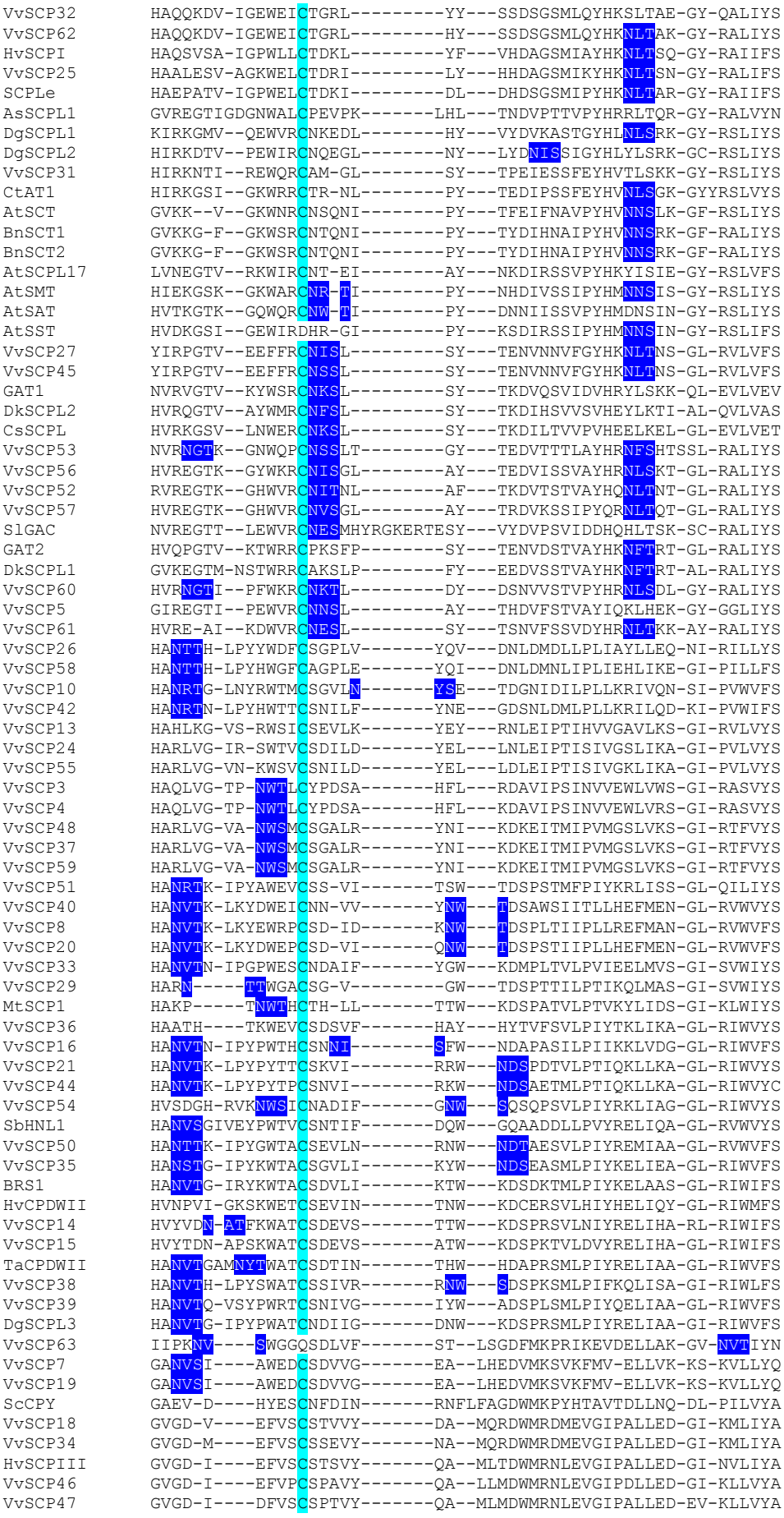
```
VvSCP32       ---PNSLPTLHLNPYSWSKVSSMIYLDSPAGVGFSFSKNTWQYN-TGDLQTASDTHEFLL
VvSCP62       ---PISLPTLHLNPYSWSKVSSMIYLDSPTGVGFSFSKNTWQYK-TGDVQTASDTHEFLL
HvSCPI        ---VKSLPKLHLNPYAWSKVSTMIYLDSPAGVGLSYSKNVSDYE-TGDLKTATDSHTFLL
VvSCP25       ---QGDLPQLHLNPYSWSKLSNIIYLDSPAGVGFSYSENLTDYR-TGDLKTASDSHAFIL
SCPLe         ---SGSLPSLHNNPYSWSKVSNIIYLDSPVGVGLSYSGNKSDYN-TGDLKTASDSHSFLL
AsSCPL1       ATGGKGFPGLLYFEDAWTKASNMIFLDAPVGAGFSYARQTEGLN-STVTGLGRHVRVFLQ
DgSCPL1       ---NGSLPTLKLNPFSWTKVSNIVFFDAPIGTGFSYSNNLNDY--PSDKKSSKQAYNVIR
DgSCPL2       ---NGSLPTLKLNPYSWTKVSNIVFLDAPPGTGFSYSRNLKGY--PSDKKSSKQGYIFIR
VvSCP31       ---HGSLPTLELNPHSWTQVSNIIFLDAPVGTGFSYATTSRASH-SGDFQATHQAHEFLR
CtAT1         ---NGSLPNLILRPHSWTKVSSIIFLDLPVSTGFTYARTEVAAQ-KSDLKLVHQAHEFLR
AtSCT         ---NGTVPPLELTSFSWTKVANILYLEAPAGSGYSYAKTRRAFE-SSDTKQMHQIDQFLR
BnSCT1        ---NGTLPPLELTSFSWTKVANILYLESPAGSGYSYAKTRRAAE-TSDTKQIHQIDQFLR
BnSCT2        ---NGTLPPLELTSFSWTKVANILYLESPAGSGYSYAKTRRAAE-TSDTKQIHQIDQFLR
AtSCPL17      ---NGSIPSLVSTTYAWTKVASILYLDQPVGTGFSYSRNPLADI-PSDTGVAKPVNEFLH
AtSMT         ---NGSAPSLFSTTYSWTKMANIIFLDQPVGSGFSYSKTPIDK--TGDISEVKRTHEFLQ
AtSAT         ---NGSNPSLFSTTYSWTKMANIIYLDQPVGSGFSYSRTPIGK--SSDTSEVKRIHEFLQ
AtSST         ---NGSVPSLVSTTYSWTKTANIIFLDQPVGSGFSYSKTPIER--TSDTSEVKKIHEFLQ
VvSCP27       ---IGGLPTFLPYPHSWTKTANIIFVDAPVGTGFSYATTSQAYT-TSDTLSVIQTLGFLR
VvSCP45       ---TGGIPTLLPYPRSWTKTANIIFVDAPVGTGFSYATTSQAYT-TSDTLSVIQTLGFLR
GAT1          ---PGGLPRLLPYKYAWTKTASILFLDAPVGTGFSYSTSADGWS-SSDTDSALETYEFLR
DkSCPL2       ---TGGLPKLRYYPDAWTKTASIIFLDQPVGTGFSYSTTQEGWP-SSDTKSSEQSYEFLK
CsSCPL        ---TGGLPKLRYYPYAWTKTASIIFLDAPVGTGFSYSRTADGWP-TSDSKSAEQSYQFLR
VvSCP53       ---DGGVPSLNLSADAWTQGLNMIYIDAPVGTGFSYSNTSQGYY-VDDAENAAQTYEFLR
VvSCP56       ---NGGLPTLELKEYTWTETLNIIYLDAPVGTGFSYSTTQEGYT-TDDYKSAAQIYEFLK
VvSCP52       ---EGGLPNLYLKENTWTKTLNIIFVDAPVGSGFSYSKTQEGYI-MEDLKYAAQTYEFLK
VvSCP57       ---TGGLPSLLLNEYTWTSGLNIIYVDTPVGAGFSYSRTQEGYY-SDDYKSSTHTYEFLN
SlGAC         ---SGNFPKLELNSYSWTKVANIIFIDQPAGTGYSYANTSEAYN-CNDTLSVTLTYDFLR
GAT2          ---DGILPTLKLNPYSWTKVASIIFIDAPVGTGFSYAETSYGYN-VSDTSSAAQTYQFLR
DkSCPL1       ---PGTLPSLRASKNSWTKVANIIFIDQPVGTGFSYGTTAAAYN-SSDTVAAAQVYKFLR
VvSCP60       ---NGSLPSLELNPYSWTKVASIIFLDAPVGTGFSYATNPDDYY-ASDTISARDNYIFIR
VvSCP5        ---SGDIPALLSNPYSWTKVASIIFLDSPVGSGFSYAQSSEGYR-TSDSLAAAHGYDFLK
VvSCP61       ---NENLPTFKLNPYSWTKLASMIFLDAPVGTGFSYSRTAEGYN-MNDTLSASQIYAFLR
VvSCP26       --------ILVKNKHSWNIESNMLYVESPIGVGFSYSNTSSNYF-WNDTRTAEDNLRFIV
VvSCP58       --------ILVKNKHSWNLESNMLYVESPIGVGFSYSNTSSDYF-WNDTRTAEDNLRFVI
VvSCP10       --------GLRRNSKSWNKASNLLFVESPAGVGWSYSNTTSDYT-CGDASTARDMRVFMM
VvSCP42       --------GVRRNSKSWNKVSNLLFVESPAGVGWSYSNTSADYN-CGDASTASDMLTFML
VvSCP13       --------ILVNNDYSWNKVANMLYLESPAGVGFSYSANTSFYAFVNDEMTARDNLKFLQ
VvSCP24       --------FLLRNEYSWNREANMLYLETPVGVGFSYSSDT-PYVTVDDKITARDNLAFLQ
VvSCP55       --------LLVRNEYSWNREANMLYLETPIGVGFSYSTDSSSYAAVNDKITARDNLVFLQ
VvSCP3        --------TLFRNKHSWNREANMLYVESPAGVGFSYSRNKSFYDDINDEVTARDNLAFLE
VvSCP4        --------TLLRNKHSWNREANMLYVESPAGVGFSYSRNKSFYDDINDEVTARDNLAFLE
VvSCP48       ------------------------------------------------------RDNLVFLK
VvSCP37       --------ILLTNKYSWNREANMLYPESPAGTGFSYSANTSFYTNLNDEITARDNLVFLK
VvSCP59       --------ILLTNKYSWNREANMLYPESPAGTGFSYSANTSFYTNLNDEITARDNLVFLK
VvSCP51       --------TLYLRPYAWNKVANTLFLESPVGVGFSYSNNSFEYNENGDKRTAQDTYAFLI
VvSCP40       --------TLYRNRFAWNKAANVLFLETPSGVGFSYSNISYNY--RGDRKTAGANYAFLV
VvSCP8        --------TLYRNQYAWNKVANVLFLESPAGVGFSYSNTTSDNQSGGDRKTANENYAFLV
VvSCP20       --------TLYRNQYAWNKVANVLFLESPAGVGFSYSNTTSDYRNGGDRKTAKDNYAFLV
VvSCP33       --------TLSYNEYAWSNVANILFLESPAGVGFSYSNTTSDYDKSGDKQTAEDNYTFLL
VvSCP29       --------TLFRNEYAWNNVSNVIFLESPAGVGFSYSNTSSDYVNVGDKKTAEDSYTFLI
MtSCP1        --------TLYRNPYAWNEVANVLFLESPAGVGFSYSNTSSDYDNSGDKSTAKDAYVFLI
VvSCP36       --------GLHFNDFAWNKEANLLFVESPVGVGFSYTNTSSDLTKLTDGFVAEDAYNFLV
VvSCP16       --------KLKFNPHTWNKAANLLFVESPVGVGFSYTNTSSDIDQLGDTITAKDSYAFLI
VvSCP21       --------KLILNDFSWNKVANILFLEAPVGVGFSYTNKSSDLLKLGDRITAEDSHAFLV
VvSCP44       --------QLILNDFSWNKVANILFLEAPVGVGFSYTNKSTDLLKLGDRITAEDSHAFLV
VvSCP54       --------GLKFNPYSWNREANMLFLESPVGVGFSYSNTTSDYEKLGDDFTANDNYAFLH
SbHNL1        --------SLLLNEYAWNKAANILFAESPAGVVFSYSNTSSDLS-MGDDKMAQDTYTFLV
VvSCP50       --------GLYLNKFSWNTLANLLFLETPAGVGFSYSNKSSDLLDTGDRRTAKDSLVFLV
VvSCP35       --------SLYLNKYSWNRVANILFLESPAGVGFSYTNTSSDLKNSGDRRTAQDALIFLV
BRS1          --------NLYLNKFAWNKDANLLFLESPAGVGYSYTNTSSDLKDSGDERTAQDNLIFLI
HvCPDWII       --------GVHMNPYSWNQVANILFLDSPVGVGYSYSNTSADILSNGDERTAKDSLVFLT
VvSCP14       --------TLYLNPYSWNQAANILFLDFPVGVGFSYSNSSFDISSNGDLRTAKDSLKFLL
VvSCP15       --------TLYLNPYSWNQVANILFLDSPVGVGFSYSNTSSDVSTNGDIRTAKDSLAFLL
TaCPDWII      --------GLVLNEYRWNKVANVLFLDSPAGVGFSYTNTSSDIYTSGDNRTAHDSYAFLA
VvSCP38       --------TLHLNPYAWNAEANLLFLDSPAGVGFSYSNTSSDLPNVGDKRTAKDAYKFLI
VvSCP39       --------TLFINPYAWNNLANLLFLESPAGVGFSYSNTTSDLYTAGDQRTAEDAYTFLI
DgSCPL3       --------TLFSNPYSWNKLANLLFLDSPAGVGYSYTNTTSDLYTTGDQRTAEDSYAFLL
VvSCP63       ---------LKPRNSTWLRKADLLFVDNPVGTGYSFVEDTKLLV-KTDVEAAVDLTTLLK
VvSCP7        --------QLEPNLGAWNRIFGLLFLDNPVGTGFSVASSPKEIP-TDQYSVAKHLFFAIR
VvSCP19       --------QLEPNLGAWNRIFGLLFLDNPIGTGFSIASSPKEIP-TDQYSVAKHLFFAIR
ScCPY         --------KPIGNPYSWNSNATVIFLDQPVNVGFSYSGSSGVS---NTVAAGKDVYNFLE
VvSCP18       --------TLSWNQYGWDQVSNILFVDQPTGTGFSYSSDEGDIR-HNEEAVSNDLYDFMQ
VvSCP34       --------TLTWNDYGWDQASNILFVDQPTGTGFSYSSSESDIR-HSEEGVSNDLYDFMQ
HvSCPIII      --------SLVWNKFGWDKISNIIFVDQPTGTGFSYSSDDRDTR-HDETGVSNDLYDFLQ
VvSCP46       --------SLMWNDWGWDKISNLLYVDQPIGTGFSYSSDLRDIR-HNEEAISNDLFDFLQ
VvSCP47       --------SLLWNEFGWDKVSNLLYVDQPIGTGFSYSSDKHDIR-HNEEGVSNDLYDFLQ
```

```
VvSCP32        RWFKEFPEFIT--NPFYVSGESYAGVYVPTLSAAIVKGIKS-GAKP--TINFKGYLVGNG
VvSCP62        RWFKEFPEFIT--NPFYVSGESYAGVYVPTLSAAIVKGIKS-GAKP--TINFKGYLVGNG
HvSCPI         KWFQLYPEFLS--NPFYIAGESYAGVYVPTLSHEVVKGIQG-GAKP--TINFKGYMVGNG
VvSCP25        KWFELYPEFLS--NPFYIAGESYAGVYVPTLAYEVVKGIKG-GIKP--ILNFKGYMVGNG
SCPLe          KWFEIYPEFLK--NPFYISGESYAGIYVPTLASEVIKGIDA-GVRP--AINFMGYMVGNG
AsSCPL1        KWMAQHPELAS--NPLYIGGDSFSGYTVTVSALEVANHPAA---SS--ELNLKGYMVGNA
DgSCPL1        KWLIDHPKFLS--NAMYVGGDSYSGKTVPITVQHISNGIEA-GDKL--LINFKGYLLGNP
DgSCPL2        KWLVDHQKFLS--NPLYISGDSYSGITIPPIGQYIFDGIES-GDEL--LINFKGYLLGNP
VvSCP31        KWLIDHPEFLS--NPVYVGGDSYSGITVPVVVQHISNGNED-DTEP--FINLKGYLLGNP
CtAT1          KWLIDHPKFLS--NEVYIGGDSYSGITVPAIVQEISQGNEK-GIQP--SINLQGYILGNA
AtSCT          SWFVKHPEFIS--NPFYVGGDSYSGKIVPGAVQQILLGNEK-GLTP--LINIQGYVLGNP
BnSCT1         SWFVDHPEFIS--NSFYVGGDSYSGKIVPGVVQQISLGNEK-GLTP--LINIKGYVLGNP
BnSCT2         SWFVDHPEFIS--NSFYVGGDSYSGKIVPGVVQQISLGNEK-GLTP--LINIQGYVLGNP
AtSCPL17       KWLDKHPEFLS--NPLYVAGNSYSGIVIPTIVQEISNGNHL-DSKP--QINLQGFVLGNP
AtSMT          KWLSRHPQYFS--NPLYVVGDSYSGMIVPALVQEISQGNYI-CCEP--PINLQGYMLGNP
AtSAT          KWLSKHPQFFS--NPFYVTGDSYSGMIVPALVQEISKGNYI-CCKH--LINLQGYVLGNP
AtSST          KWLIKHPQFLS--NPFYVVGDSYSGMIVPALVHEISKGNYI-CCNP--PINLQGYVLGNP
VvSCP27        NWLNDHPDFKL--NPLFIGADSYSGLIAPIIAQEIMDGNGV-GEEP--HINLKGYLIGSP
VvSCP45        NWLNDHPDFKL--NPLFIGADSYSGLIAPIIAQEIMDGNGV-GEEP--HINLKGYLIGSP
GAT1           KWLIEHPKYLP--LQLYVGGDSYSGIIVPLVVKHIVDAIDE-HTVP--RFNLQGYLVGSP
DkSCPL2        KWLEENPQYLK--VQLFVGGDSYAGKIVPLVTRLIADGNKN-GGTP--YLNLKGMVLGSP
CsSCPL         EWFDEHPQYLA--VQLFVGGNSYSGMIVPLVTKKIVDGNKD-GVKP--FMNLKGYLLGSP
VvSCP53        KWLVQHPDFLG--NELYIAGVSYSGIPVPMIVNEIIEGNIL-GLSP--GMNIKGYVLGSP
VvSCP56        KWLIQHPEFLK--NNLYIGGDSYSGIPVPMIVQDIYYDSER-GGSPRLNLNLQGYVLGNP
VvSCP52        KWLVDHPEFLK--NELYVGGDSYSGIPVPMVVQEIYYGSP--------SLNLQGYVLGNP
VvSCP57        KWLLDHPEFLK--NNLYVGGDSYSGIVLPMITEKIYYGNGI-GT--FLQMNLQGYILGNP
SlGAC          KWLMDHPEYLN--NPLYVGGDSYSGIFVALLTRKIYDGIEV-GDRP--RVNIKGYIQGNA
GAT2           KWLTFHPNFAG--NPLYIGGDSYSGIVAPILIKDILHGLEV-GLQP--KIELQGYLLGNP
DkSCPL1        KWLMYNPKFGA--NPLYVGGDLYSGITVPLLVQTILDGIGS-GSLP--RMQLKGYLLGNP
VvSCP60        KWLIDHPRFLY--NPLYIGGDSYSGIIVPILTLEIANGIQM-GLKP--LMTLMGYILGNP
VvSCP5         KWLIDHPEFLR--NRLYIAGDSYSGLFVPIIAQKISDGNEA-GQEP--HMNLNGYLLGNA
VvSCP61        KWLINHPKFQK--NPLYVSGDSYSGIIIPMVVQEISNGNDE-GKEP--KMNIQGYTIGNP
VvSCP26        NWFEEFPYYKD--SELFLTGESYAGHYIPQLAALLVEYNKR-PNIR--PIKLKAIALGNP
VvSCP58        NWLEEFPNYKD--SELFLTGESYAGHYIPQLAALIVEYNQK-PNIR--PIKLKSIALGNP
VvSCP10        KWLEKFPAFKS--RALFLTGESYAGHYIPQLAVALLDYNSH-STGF--KFNLKGVAIGNP
VvSCP42        KWFKKFPGYKL--RPLFLTGESYAGHYIPQLANVLLDYNKK-SKDF--KFNIKGVAIGNP
VvSCP13        RWFLKFPEYKN--RDLFLTGESYAGHYVPQLAQLIVQS------KV--KFNLKGVAIGNP
VvSCP24        RWFLKFPQYKH--RDLFITGESYAGHYVPQLAELMIRFNKK---EK--LFNLKGIALGNP
VvSCP55        KWLLKFPQYKN--RDLFITGESYAGHYVPQLAELMLQFNKK---EK--LFNLKGIALGNP
VvSCP3         GWFMKFPKYRN--RELFITGESYAGHYVPQLAQLVINS------GK--NFNLKGILIGNP
VvSCP4         GWFMKFPKYRN--RELFITGESYAGHYVPQLAQLVINS------GK--NFNLKGILIGNP
VvSCP48        NWFIKFPQYKN--SELFIAGESYAGHFVPQLAQLILES------RV--KFNLKGILMGNP
VvSCP37        NWFIKFPQYKN--SELFIAGESYAGHFVPQLAQLILES------SV--KFNLKGILMGNP
VvSCP59        NWFIKFPQYKN--SELFIAGESYAGHFVPQLAQLILES------RV--KFNLKGILMGDP
VvSCP51        NWFRRFPHYKN--RDFYIMGESYAGFYIPELADTIIRRNMKAVSSS--IIHLKGIMIGNG
VvSCP40        NWLERFPEYKK--RDFYIAGESYAGHFVPQLAHVILHHNKK-ANRT--IINLKGITIGNA
VvSCP8         NWLERFPEYKK--RDFYISGESYAGHYVPQLAHTILHHNKK-ANGP--IINLKGIIIGNA
VvSCP20        NWLERFPEYKK--RDFYISGESYAGHYVPQLAHTILHHNKK-ADGP--IINLKGIIIGNA
VvSCP33        NWLERFPEYKT--RDFFITGESYAGHYVPQLSQKILQNNKI-TNQT--VINLKGIAIGNA
VvSCP29        NWLERFPQYKT--RDFFITGESYSGHYVPQLAYTILSNNNK-TNQT--VINLKGIAIGNA
MtSCP1         NWLERFPQYKT--RDFYITGESYAGHYVPQLASTILYNNKL-YNNT--IINLKGISIGNA
VvSCP36        NWLKRFPQYKA--HDFFISGESYAGHYVPQLAELVYDRNKDRTKYP--LINLKGFIVGNP
VvSCP16        HWFKRFPQFKP--HDFYIAGESYAGHYVPQLAEVIYDHNKHVSKKL--HINLKGFMIGNA
VvSCP21        QWFKRFPSLKT--HDFYITGESYAGHYVPQLAELIYERNKRSSKDF--YINLKGFMIGNA
VvSCP44        QWFKRFPSFKS--HDFYITGESYAGHYVPQLAELIYERNRKSTKDS--YINLKGFMIGNA
VvSCP54        KWFLKFPSYRK--RIFYIAGESYAGKYVPELAEVIYDKNKD-P-SL--FIDLRGILLGNP
SbHNL1         KWFERFPHYNY--REFYIAGES--GHFIPQLSQVVYRNR---NNSP--FINFQGLLVSSG
VvSCP50        RWLERFPRYKH--REVYINGESYAGHYVPQLAREIMAYNAK-YKH---AINLKGIMVGNA
VvSCP35        RWMSRFPKYKH--REFYIAGESYAGHYVPQLAKKIHDYNKA-SSHP--IINLKGFMVGNA
BRS1           KWLSRFPQYKY--RDFYIAGESYAGHYVPQLAKKINDYNKA-FSKP--IINLKGFLVGNA
HvCPDWII       KWLERFPQYKE--REFYLTGESYAGHYVPQLAQAIKRHHEA-TGDK--SINLKGYMVGNA
VvSCP14        EWFERFPQYKG--RDFYITGESYAGHYVPQLSQAIVRYNFA-TKAK--SINLKGYMVGNA
VvSCP15        KWFERFPQYKG--RDFYITGESYAGHYVPQLSQAIVRHNSA-TKAE--SINLKGYMVGNA
TaCPDWII       KWFERFPHYKY--RDFYIAGESYAGHYVPELSQLVHR-----SKNP--VINLKGFMVGNG
VvSCP38        NWLQRFPQYKH--RPFYIAGESYAGHYIPELSQIIVQRNKG-MKNP--AINFKGFLLGNP
VvSCP39        NWFERFPQYKH--RDFYIAGESYAGHYVPQLSQIVYRRNKG-IQNP--VVNFKGFLVGNA
DgSCPL3        NWLERFPRYKH--REFYIVGESYAGHYIPQLSQIIYRRNMG-VQNP--VLNFKGFMVGNA
VvSCP63        EIFNRNESLQQ--SPLYIVAESYGGKFAVTLGLAALEAI--EAGNL--KLKLGGVALGDS
VvSCP7         SFIELDPLFKS--RSIYITGESYAGKYVPAIGYHILKKNSRLPESQ--RVNLRGVAIGNG
VvSCP19        SFIELDPLFKS--RSIYITGESYAGKYVPAIGYYILKKNAQLSESQ--GVNLRGVAIGNG
ScCPY          LFFDQFPEYVNKGQDFHIAGESYAGHYIPVFASEILSHK-----DR--NFNLTSVLIGNG
VvSCP18        AFFAKHPEFVK--NDFYITGESYAGHYIPAFAARVQKGNK-ANEGV--HINLKGFAIGNG
VvSCP34        AFFKKHPEFVR--NDFFITGESYAGHYIPAFAARVQKGNK-AKEGV--HINLKGFAIGNG
HvSCPIII       VFFKKHPEFIK--NDFFITGESYAGHYIPAFASRVHQGNK-KNEGT--HINLKGFAIGNG
VvSCP46        AFFEEHPLFTN--NDFYIAGESYAGHYIPALAARIHRGNK-AKGGI--HINLKGFAIGNG
VvSCP47        AFFEEHPQFAD--NDFYITGESYAGHYIPAFAARVHRGNK-AKEGI--HIKLKGFAIGNG
```


VvSCP32 VTDMEFDANALVPFTHGMG----LISSEMFEKARDNCGG----NYY--------SNESKS
VvSCP62 VTDMEFDANALVPFTHGMG----LISSEMFEKARDNCGG----NYY--------SNESKS
HvSCPI VCDTIFDGNALVPFAHGMG----LISDEIYQQASTSCHG----NYW--------NATDGK
VvSCP25 VTDEEFDGNALVPFAHGMG----LISDELFQDISNLCQG----NYY--------NSLDEN
SCPLe VADDIIDGNAIVPFQHGMG----LISDDLYEEAVVACHG----NFY--------EPVDSN
AsSCPL1 RGEVNNDNACRIPYLHGMG----LISDELYEAALSSCVV----GTDSKNK---QQQSAAR
DgSCPL1 ITDPDLGEYSTIQFAHGMG----IISEELFESITKNCKG----EYV---N---INPDNVK
DgSCPL2 KTDLYYDHNSKIPFAHGMA----LLSTELFESVKKNCGG----EYY---N---ADPENVQ
VvSCP31 VTEQGTETTAQFRFAHGMA----LISDELYESLKTSCGD----EYP---F---KYPINIQ
CtAT1 FTTRKEE-NYAIPFAHGMA----LISDELYESLQKNCKG----EYI---D---VDTKNAL
AtSCT VTDKNIETNYRVPFAHGMG----LISDELFESLERSCGG----KFF---N---VDPSNAR
BnSCT1 AVRTNLEPNHRVSFAHRMG----LISDELHESLERNCGG----KFF---N---VDPSNAK
BnSCT2 AVRTNLEPNHRVSFAHRMG----LISDELHESLERNCGG----KFF---N---VDPSNAK
AtSCPL17 ATDTDIDLNSRIPFAHGKA----LISDEHYESLKRSCQG----NYI---S---VNPRNTK
AtSMT VTYMDFEQNFRIPYAYGMG----LISDEIYEPMKRICNG----NYY---N---VDPSNTQ
AtSAT ITYAEHEKNYRIPFSHGMS----LISDELYESLKRNCKG----NYE---N---VDPRNTK
AtSST ITHIEFEQNFRIPYAHGMS----LISDELYESLKRICKG----NYF---S---VDPSNKK
VvSCP27 HTDTTIELNSRIVYAHRMA----LISDALYEAAKTGCNG----RYV---D---IDPSNAK
VvSCP45 HTDTTIELNSRIVYAHRMA----LISDALYEAAKTGCNG----RYV---D---IDPSNAK
GAT1 TTDENINTNAKVVFAHRLA----LISDELYEAAKENCNG----NYA---D---VDPSNTK
DkSCPL2 RTDNIIDENSKVVFAHRMA----LISDEMYENAKEACNG----SYS---N---AAPNNTA
CsSCPL RTDSVIDENSKVVFAHRMA----LISDEIYENAKKGCND----TYV---N---IDPANTA
VvSCP53 VTDSFIDDNSKIPFAHGLS----LISHELYNSAKTNCEG----NYV---N---VSSEAS-
VvSCP56 VTDAYIDKNSRVPFAHRLT----LISDGLYESAKANCNG----DYV---N---ANASSEQ
VvSCP52 LTDTDNDVNSRIPFAHRLT----LISDELYESAKTSCNG----DYV---T---VNASNEQ
VvSCP57 VTDSYIDGNAQIKIAHRLT----LIPDNLYESAKASCNG----DFV---T---VNASNEE
SlGAC LTDRSIDFNGRVKYANHMG----LISDKIYQSAKANCNG----NYI---D---VDPNNIL
GAT2 LTDYLIDDNSRIPYAHRVS----LISDGLYKAAKETCNG----DYG---N---VDINNTL
DkSCPL1 LTDDFIDANSKIPYAFRVN----LLSDELYEDAEVSCNG----DFV---N---VDFNNTN
VvSCP60 VTHLHNDENSRIPFAHRIA----LISDELYESAKNACKG----EFI---D---PDESNGE
VvSCP5 LVDENIDFNSRVPFAHRMT----FLSDKLYKKTEASCNG----KYL---K---ADPSNGQ
VvSCP61 VTDHFSDFNSRIEYTHRVG----ILSDELYEELKESCNG----KYV---Y---VDPSNVE
VvSCP26 LLDLDISVL-AGDYLWSHG----AISDDTLLLEKTVCND----SKYLREYY--HGQLSKE
VvSCP58 LLDLDISVL-AADYLWAHG----AISDHTLMLEKTVCNY----SKFLREYI--HGQLSEG
VvSCP10 LLRLDRDSAATYEFFWSHG----MISDEIGLTITKECDF----DDYV---YASPHNVSFS
VvSCP42 LLQLDRDVPAVYEFFWSHG----MISDEVGLAIMNDCNF----EDYT---FSGTHNVSTE
VvSCP13 LLEFNTDFNSRAEYMWSHG----LISDITYEAFTVICNY----SQVRREIV--MGSLSPA
VvSCP24 VLEFATDLNSRAEYFWSHG----LISDSTYRLFTSACNY----SRYVSEYY--RDSVSSV
VvSCP55 VLEFATDLNSRAEFFWSHG----LISDSTYKMFTSFCNY----SRYVSEYY--RGSVSSI
VvSCP3 LLEFDTDINAQGDFFWSHG----LISDSTHALLTSTCNY----SQIMRWVYNISESLSPE
VvSCP4 LLEFDTDMNAQGDFFWSHG----LISDSTHALLTSTCNY----SQIMRWVYNISESLSPE
VvSCP48 LMDFDTNYNSVPHFYWSHG----LISDSTYNLFSSKCNY----SRMNRE--QTSGSLSPA
VvSCP37 LMDFDTNYNSVPHFYWSHG----LISDSTYNLFSSKCNY----SRMNRE--QTSGSLSPA
VvSCP59 LMDFDTNYNSVPHFYWSHG----LISDSTYNLFSSKCNY----SRMNRE--QTSGSLSPA
VvSCP51 IMNDMTDNRGFYDYLWSHA----LISDKTHQGLVEYCKF----PD----------SYE
VvSCP40 AIHDETDWLGMYQYFGSHA----LVSPRTTRQIEKHCDF----SPGV-------TNQNKE
VvSCP8 VIDDEADDIGRYQYLGSHA----LVSEKTIHQMEKHCNF----SPGA-------TSQSKE
VvSCP20 VINDETDELGMYQYFGSHA----LVSEKTIRQMEKHCNF----SPGA-------ASQSKE
VvSCP33 WIDYETGLKGMYDFFWTHS----LISDEINEGINLNCNF----SSET--------TISDV
VvSCP29 WIDDNTSLKGIYDYIWTHA----LSSDESNAGIQKYCDF----TTG---------NFSTK
MtSCP1 WIDDATNLKGIYDNLWTHA----LNSDQTHELIEKYCDF----TKE---------NVSAI
VvSCP36 ETNDYYDYKGLLEYAWSHA----VISDQLYYKSKQVCDF----KVAD---------WSSE
VvSCP16 LLDDDTDQRGMVSYAWDHA----VISDRVFFDIKKACNF----SAEP---------VTEE
VvSCP21 VINDETDDMGLIEFAWSHA----IISDQIYHGIMKNCDF----KSGN---------LTNL
VvSCP44 VINDETDDMGLIEFAWSHA----IISDQLYHGIIKECDF----IRDN---------PTNL
VvSCP54 ETCDADDWRGLVDYAWSHA----VVSDETHKIIRENCDF----YSED-------PWSNDN
SbHNL1 LTNDHEDMIGMFELWWHHG----LISDETRDSGLKVCPG----TSFM-------H-PTPE
VvSCP50 VTDNYYDNLGTVTYWWSHA----MISDKTYRQLINTCDF----HRQ--------K-ESNE
VvSCP35 VTDNYYDSIGTVAFWWSHS----MISDRSYRSIMDHCDF----IAE--------R-TSEK
BRS1 VTDNQYDSIGTVTYWWTHA----IISDKSYKSILKYCNF----TVE--------R-VSDD
HvCPDWII LTDDFHDHYGIFQYMWTTG----LISDQTYKLLNIFCDF----ESFV-------H-TSPQ
VvSCP14 LTDDFHDHLGLFQFMWSVG----MISDQTYKLLNVFCDS----QSFI-------L-SSEL
VvSCP15 LTDDFHDHLGVFQFMWSAG----MISDQTYKLLNVFCDF----QPFI-------H-SSAS
TaCPDWII LIDDYHDYVGTFEFWWNHG----IVSDDTYRRLKEACLH----DSFI-------H-PSPA
VvSCP38 LIDDYYDNKGTHEFWWSHG----LISDSTYEALKEACAN----DTFL-------F-PKDK
VvSCP39 VTDDYHDYIGTFEYWWTHG----LISDSTYKILRVACDL----GSSM-------H-PSSE
DgSCPL3 VIDDFHDYVGTFEYWWTHG----LISDATYQSLNANCDF----EVSE-------H-PSEA
VvSCP63 WISPEDFVFSWGPLLKDVS----RIDNQGLQKSNSLARKIRQQLIDGQYV---DATSSWS
VvSCP7 LTDPVRQVATHAASAYFSG----LINGKQKTQLEKAQLE-AVKLIK---------EGNWS
VvSCP19 LTDPVRQVATHAASAYFSG----LINGKQKTQLEKAQLE-AVKLIK---------EGNWS
ScCPY LTDPLTQYNYYEPMACGEGGEPSVLPSEECSAMEDSLER-CLGLIES-CY---DSQSVWS
VvSCP18 LTDPSIQYKAYTDYALNMK----IIGKSDYDSINELIPE-CEESAKS-CG---PD-GGDA
VvSCP34 LTDPAIQYKAYTDYALTMK----IIGKSDYDSINELIPD-CEESAKS-CG---PA-GGVA
HvSCPIII LTDPAIQYKAYTDYALEMN----LIQKADYERINKFIPP-CEFAIKL-CG---TN-GKAS
VvSCP46 LTNPQIQYKAYTDYALEMG----MIEKTDYDRINKVLPV-CEMAIKL-CG---TD-GTIS
VvSCP47 LTDPQIQYKAYTDYALDMG----IIQKPDYDRINKVLPV-CEMAIRL-CG---TD-GTIS


VvSCP32 CIEELNKI-YNAISGLNKYN-----ILEPCYHRPAKKGEETG---NTTLPLSFKQLGATN
VvSCP62 CIEELNKI-YNAISGLNQYD-----ILEPCYHRPTKKGEETG---NTTLPLSFKQLGATN
HvSCPI CDTAISKI-ESLISGLNIYD-----ILEPCYHSRSIKEVNLQ---NSKLPQSFKDLGTTN
VvSCP25 CESKLSKV-DKDIEGLNIYD-----ILEPCYHEKSPET-SLG---NIRLPSSFQKLGETD
SCPLe CSEKLNKI-DQVVYDLNVYD-----ILEPCYHSKKPSVITTG---NSRLPMSFRKLGETE
AsSCPL1 CSEAQQAI-SEATTDLNPAH-----ILEPACGADFSPRAPYLSLTTPSSSSSS----SSS
DgSCPL1 CTKDIQRV-SKCTNGLNRAH-----ILEPHCLLVTPKP-------NEII--------KER
DgSCPL2 CRKDLQLV-SECTKGVNDPH-----ILEPKCLLVTPKS-------NKWN--------KRR
VvSCP31 CIKDVQAF-YKCISGIQFGQ-----ILEPVCGFGSLKPE------DIFL--------SGR
CtAT1 CSRVMESY-NEVISGISFSH-----ILEPNCDWVDTETS-------------------LR
AtSCT CSNNLQAY-DHCMSEIYSEH-----ILLRNCKVDYVLAD----TPNIRT--------DRR
BnSCT1 CSNGLLAY-HQCISEIYIEQ-----ILLPNCKVDYVLADISQTLPNIRT--------SRR
BnSCT2 CSNGLLAY-HQCISEIYIEQ-----ILLPNCKVDYVLADISQTLPNIRT--------SRR
AtSCPL17 CLKLLEDF-KKCVSGISEEY-----ILKPDCMWL--------------------------
AtSMT CLKLTEEY-HKCTAKINIHH-----ILTPDCDVT--------------------------
AtSAT CVRLVEEY-HKCTDKINTQH-----ILIPDCDKKG-------------------------
AtSST CLKLVEEY-HKCTDNINSHH-----TLIANCDDSN-------------------------
VvSCP27 CVEALESI-SLCIEQVSLQD-----ILEPKCSFISPKQNKEI-----------------R
VvSCP45 CVEALESI-SLCIEQVSLQD-----ILEPKCSFISPKQNKEI-----------------R
GAT1 CLSSLGEI-QHCVKDLFRND-----ILEPKCVFESPEPT--------------------R
DkSCPL2 CHLAIEEI-TRCIRDLFRGN-----ILEPRCLFVAPPQTESDAIHAERRSLQAQA--QEE
CsSCPL CIVALGNI-KTCIKDLYRND-----ILEPKCVFATPDPGEEPAA---------------R
VvSCP53 -----LNF-HQLLRYINVAQ-----VLHPYCYPFTVKPSERQGN-------------RRS
VvSCP56 CESDVQEI-EELLRDINIQQ-----ILDPDCTFSSPIPNEEKSN-------------LQR
VvSCP52 CVADMEAI-SKLIDQIYIMQ-----VLEPNCGISSRKPKEGE------------------
VvSCP57 CVADMEAI-SELISPIYTMQ-----VLEPNCGISSQKPNKWKS--------------QQR
SlGAC CLNDLQKV-TRCLKNIRRAQ-----ILEPYCDLPYLM-------------------------
GAT2 CVEALQTI-KMCLLQINIAM-----ILEPQCAFA--SPQTTELQ-------------WDL
DkSCPL1 CVAVLQGI-KENLQLLNEAQ-----NFGPLCALA--KPKGEGIQ-------------WGA
VvSCP60 CMEVLAVI-TKCTEKLNSAH-----ILEPVCALDSPKPKESKLK-------------WSL
VvSCP5 CTENLKVV-NKCMEKINLPH-----VLEPKCGRPL-SWKPNALK-------------WES
VvSCP61 CTNNLKVY-TQCINKIYSAH-----ILEPSCAM-L-SPNPNASK-------------LDR
VvSCP26 CKDVFNRVLDEISGDVEKGD-----LLMPKCLSSNSAQQFRLKG-------------LQG
VvSCP58 CNNVYNRVVNEIGNDVRQDD-----LLLPICLSSNSAQQFKLKG-------------QRG
VvSCP10 CNQALSEANSIVGEYINNYD-----VILDVCYPAIVEQELRLRR-------------MAT
VvSCP42 CSTALNDAYSIVGSYINPYD-----VILDVCYPSIVQQELRLRK-------------VVT
VvSCP13 CSGVISQVSRELGKHIDSYD-----VTLDVCLPSVVSQSERLNQ---------------P
VvSCP24 CSRVMAQVSRETSKFVDKYD-----VTLDVCLSSVLSQSKVISP---------------Q
VvSCP55 CSRVMSQVGRETSRFVDKYD-----VTLDVCISSVLSQSKVLSP---------------Q
VvSCP3 CYEVYNKSAGEIGGSVDPFD-----VLGDICLSS--------------------------
VvSCP4 CYEVYNKSAGEIGGSVDPFD-----VLGDKCLSS--------------------------
VvSCP48 CLAVRSQYSQEVGDSVDRFD-----VTLNSCLPSVDP------Q---------------P
VvSCP37 CLAVRSQYSQEVGDSVDRFD-----VTLNSCLPSVDP------Q---------------P
VvSCP59 CLAVRSQYSQEVGDSVDRFD-----VTLNSCLPSVDP------Q---------------P
VvSCP51 CKKLEDHI-ELEVGLIDFYN-----IYAPVCLRASNSS----------------------
VvSCP40 CNAAFEEV-DPNIANIGIYN-----IYGPVCLDTNLTA----------------------
VvSCP8 CTEAVDEV-HSNIDVIDIYN-----IYSPLCFNTILTA----------------------
VvSCP20 CTKASDEV-DDNIDVIDIYN-----IYAPLCFNTNLTV----------------------
VvSCP33 CEQYLDAA-DAAVGYIYIYD-----IYAPLCSSSSNST----------------------
VvSCP29 CLDYTYQA-EGEVGNIDIYN-----IYAPLCHSSGPTS----------------------
MtSCP1 CNNATDKA-FVETGKIDIYN-----IHAPLCHDSSLKN----------------------
VvSCP36 CITNMNKV-FDDYREIDIYN-----IYAPSCLLNTTSSSAELNG-------------NGF
VvSCP16 CNIALGKY-FEVYEIIDMYS-----LYAPTCEDDATSSTTSFVA-------------RQL
VvSCP21 CIKYVEGF-FEAYLDIDVYS-----IYTPVCLSSSKETYRKL-----------------
VvSCP44 CSNHIKGL-LEAYSDIDMYS-----IYTPVCLSSSKETYRKF-----------------
VvSCP54 CSDAVGEV-LDQYKRIDIYS-----LYTSVCTKTSKRSDDNSMQ-------------VLF
SbHNL1 CTEVWNKA-LAEQGNINPYT-----IYTPTCDREPSPYQRRFWA-------------P--
VvSCP50 CESLYSYAMDQEFGNIDQYN-----IYAPPCNNSDGSGATRQTI-------------RLP
VvSCP35 CDEAVSYAINHEFGDIDQYS-----IYTPSCMALPNSSTIR--S-------------PRF
BRS1 CDNAVNYAMNHEFGDIDQYS-----IYTPTCVAAQQKKNTTGFF-------------VRM
HvCPDWII CDKILDIA-STEAGNIDSYS-----IFTPTCHSSFASSRNKVVK-------------R--
VvSCP14 CDKIMDIA-REEIGNIDLYS-----IFTPPCSVKIGF-SNQLMK-------------K--
VvSCP15 CDKIMDIA-SEEMGNVDPYS-----IFTPPCSVKVGF-SNQLMK-------------R--
TaCPDWII CDAATDVA-TAEQGNIDMYS-----LYTPVCNITSSSSSSSSL-------------SQQ
VvSCP38 CNNALTGA-YKEFGDIDPYN-----IYSGPCREVA----------------------TLG
VvSCP39 CTKALNLA-EAEQGNIDPYS-----IFTRPCNDTS----------------------SLR
DgSCPL3 CDKTMDIA-YVEQGNIDAYS-----IFTPTCSATI----------------------SQR
VvSCP63 ELEGVISR---SSNSVDFYNFLLDSNMDPLSLTSVELR----------------------
VvSCP7 EA--TNAR---N--------RVLNMLQD--------------------------------
VvSCP19 EA--TNAR---N--------RVLNMLQD--------------------------------
ScCPY CVPATIYC---N--------NAQLAPYQ--------------------------------
VvSCP18 CETAYTNC---N--------FIFNSILN--------------------------------
VvSCP34 CDTAYYSC---N--------QIFQSIIN--------------------------------
HvSCPIII CMAAYMVC---N--------TIFNSIMK--------------------------------
VvSCP46 CMASYFVC---N--------SIFSKIMA--------------------------------
VvSCP47 CMASYFVC---N--------TIFSSIMA--------------------------------


VvSCP32 RPLPVRTRMFGRAWPFHAPVKDGILPLWPELMKKKTIPCT--DDQVASVWLNDKGVRTAI
VvSCP62 RPLPVRTRMFGRAWPFRAPVKDGILPLWTELIKQNPIPCT--DDQVASAWLNDKGVRTAI
HvSCPI KPFPVRTRMLGRAWPLRAPVKAGRVPSWQEVA--SGVPCM--SDEVATAWLDNAAVRSAI
VvSCP25 RPFAVRKRMFGRAWPLRAPVREGLVPTWPQLLNSGSVPCT--DDEVATSWLNNKAVREAI
SCPLe RPLPVRKRMFGRAWPYKAPVRAGHVPTWPEILNSVEVPCT--DDRVATLWLNNADVRKAI
AsSCPL1 SSSSSS-----SSYYYLSL-SSVRSRTP---TKEMLLECRVYGYELSYMWANDAEVRENL
DgSCPL1 RLLEKN--------------SPTTFALP---PVDPGFECRSYGYMLSYYWANDKEVQKAL
DgSCPL2 SSLEMT--------------LPNNLRFP---FGIPEFGCRSYGYMLSYYWANDKEVQKAL
VvSCP31 RYLIGK---------------LRERRPE---PSLSAFECRTDGYILAPYWANNATVQEAL
CtAT1 RSLIQR--------------HHGKKFLN---TRLPALSCRTYANFQSSFWANDDNVRSAL
AtSCT RVMK-E----------------FSVNDS---SSLPPPSCFTYRYFLSAFWANDENVRRAL
BnSCT1 RELK-E----------------FSRNDS---SSLPPPSCFTYRYFLSAFWANDENVRRAL
BnSCT2 RELK-E----------------FSRNDS---SSLPPPSCFTYRYFLSAFWANDENVRRAL
AtSCPL17 ------------------------------------YSCMANLHSLSEYWANEKSVRKAL
AtSMT ---------------------------------NVTSPDCYYYPYHLIECWANDESVREAL
AtSAT --------------------------------HGITSPDCYYYLYFLIECWANNERVREAL
AtSST --------------------------------TQHISPDCYYYPYHLVECWANNESVREAL
VvSCP27 RSLREN------------S-RSFLLPSQ---YRTGNDWCRNFEHSLSDIWANYKSVQDAL
VvSCP45 RSLREN------------S-RSFLLPSQ---YRTGNDWCRNFEHSLSDIWANYKSVQDAL
GAT1 RSLDEK------------P-GDFIL-NT---PKLEEFWCRNFNYALSYIWANDESVQEAL
DkSCPL2 DEDEDG------------T-LDFLL-SP---PRIQNLWCRAFNYVLAYEWGNDIAVQEAL
CsSCPL RSLDEG------------P-SDFLL-SP---PMIPNLWCRNFNYVLSYIWSNDDTVQEAL
VvSCP53 SL-----------------------------EEANYRSCDLYSSVPISIWANDESVRAAL
VvSCP56 SLAENP---------------TDFLSQL---GEETMYFCHDYMYILSETWANNRDVREAL
VvSCP52 ---LNH---------------THFLTQL---GEKSAYFCHEYNYVFSEIWANNKDVREAL
VvSCP57 SLIENS---------------KHFPSGL---GKKAAYHCHEYMYVFSEIWSNDESVREAL
SlGAC GILQET------------P-TN--GQSV---FPIAGPWCREKNYIYSYVWANDKAVQKAL
GAT2 RVQENT------------T-M-NYLLSL---SRIPELRCRSFSYVLSYKWLNDINVQNAL
DkSCPL1 EEAEFT----------------DSLILQ---DIIPQLTCRSSSWMLSYIYMNDEGVQRAL
VvSCP60 NHIEDV------------S-SDMISLPV---PQRSELWCRNYNYLLSYIWENDEAVQEAL
VvSCP5 IPLEEN------------F-SDFLLSPI---RQLPEPTCRLYKFLFSYIWANDRRVQKAL
VvSCP61 SSLQEI------------N-SIGLLLSQ---PQKPEPWCRSYNYVFSYLWANDKTVQKAL
VvSCP26 KI----------------------YAEIDRRTRGTIPDPCLP---DRIFTYLNNPQVQKAL
VvSCP58 TI----------------------HAAIARRTRETIPDPCLS---DRILTYLNNPQVQKAL
VvSCP10 KM----------------------SV---------GIDVCMT---YERSFYFNLPEVQKAL
VvSCP42 KI----------------------SI---------GVDVCMT---AERTFYFNLPEVQKAL
VvSCP13 RG----------------------TE---------KIDVCVE---DETIKYLNRKDVQKAL
VvSCP24 QV----------------------AE---------TIDVCID---DKTVNYLNRKDVQKAL
VvSCP55 QV----------------------TE---------TIDVCVE---DETESYLNRRDVQKAL
VvSCP3 -----------------------------------EEVCLT---DEVDVYLNRKDVQKSL
VvSCP4 -----------------------------------EEVCLT---DEVDVYLNRKDVQKSL
VvSCP48 QV----------------------TE---------NVDVCIG---DEVNKYFNREDVQKSL
VvSCP37 QV----------------------TE---------NVDVCIG---DEVNKYFNREDVQKSL
VvSCP59 QV----------------------TE---------NVDVCIG---DEVNKYLNREDVQKSL
VvSCP51 -------------------------RKPK--RHGGFDPCEA---DYVLRYLNLPQVQEAL
VvSCP40 -------------------------KPKK-VTPLQFDPCSY---DYVHAYLNRPDVQEAF
VvSCP8 -------------------------KPKK-VTP-EFDPCSD---YYVSAYLNRADVQKAL
VvSCP20 -------------------------KPKK-VTP-EFDPCSD---YYVYAYLNRADVQKAL
VvSCP33 -------------------------R-----PISVFDPCSE---DYIQTYLNIPEVQKSM
VvSCP29 -------------------------RSV--GSVNDFDPCSD---YYVESYLNLAEVQKAL
MtSCP1 -------------------------GSSTGYVSNDFDPCSD---YYVTAYLNRPEVQKAL
VvSCP36 ----------------------------RRMRVPGGYDPCFS---IYAAEYFNRPDVKLAL
VvSCP16 PL------IRGNV----AP-KTFSKFPAWHKRPTGYDPCAS---DYTTVYLNRPEVQAAL
VvSCP21 -----------VT----AP-RLFAQHDLWHQLPSGYDPCTE---DYAEKYFNREDVQKAL
VvSCP44 -----------VT----AP-RLFTQHDLWHQLPSGYDPCTE---DYAEKYFNREDVQKAL
VvSCP54 KR----------------------TSRMMPRIMGGYDPCLD---DYAKAFYNRADVQKAL
SbHNL1 HG----------------------RAAPPPLMLPPYDPCAV---FNSINYLNLPEVQTAL
VvSCP50 HR----------------------SH-RIFRQISGYDPCTE---KYAEIYYNRPDVQKAL
VvSCP35 KN----------------------SL--VRRRVSGYDPCTE---NYAEKYYNRPDVQKAM
BRS1 KN----------------------TL-LRRRLVSGYDPCTE---SYAEKYFNRPDVQRAM
HvCPDWII -------------------------LRSVGKMGEQYDPCTE---KHSIVYFNLHEVQKAL
VvSCP14 -------------------------LIMASGISRKYDPCTE---QHSAVYYNLPEVQQAL
VvSCP15 -------------------------LIRVGRISERYDPCTE---QHSVVYYNLPEVQQAL
TaCPDWII RR----------------------SRGRYPWLTGSYDPCTE---RYSTAYYNRRDVQMAL
VvSCP38 NN----------------------SKLPLPWTFRGNDECIV---RYTRKYMNRGEVQKAF
VvSCP39 RN----------------------LRGHYPWMSRAYDPCTE---RYSEVYFNLPEVQTAL
DgSCPL3 RN----------------------LRKIRPRISVPYDPCTE---KHSTVYFNLPEVQKAL
VvSCP63 ----EQFAKRRYLRYLDS-----------LRLSPGGDVDID-T------LMNGIIKDKLR
VvSCP7 --------MTGLATLYDL-----------TRKV-----PYE--LELVGEFLSSEGVKKAL
VvSCP19 --------MTGLATLYDL-----------TRKV-----PYE--FELVGEFLSSEGVKKAL
ScCPY ---------RTGRNVYDI-----------RKDCEGGNLCYP-TLQDIDDYLNQDYVKEAV
VvSCP18 --------VAGNINYYDI-----------RKQCEG-SLCYD--FSNLESFMGLKSVKKAL
VvSCP34 --------VAGNINYYDI-----------RKQCEG-SLCYD--FSNLENFMGLKSVKEAI
HvSCPIII --------LVGTKNYYDV-----------RKECEG-KLCYD--FSNLEKFFGDKAVRQAI
VvSCP46 --------LSGDTNYYDI-----------RKTCEG-SLCYD--FSNMEKFLNQKPVRAAL
VvSCP47 --------IAGDANYYDI-----------RKKCEG-SLCYD--FSNMERFLNQRSVRDAL


VvSCP32 HAQQKDV-IGEWEICTGRL---------YY---SSDSGSMLQYHKSLTAE-GY-QALIYS
VvSCP62 HAQQKDV-IGEWEICTGRL---------HY---SSDSGSMLQYHKNLTAK-GY-RALIYS
HvSCPI HAQSVSA-IGPWLLCTDKL---------YF---VHDAGSMIAYHKNLTSQ-GY-RAIIFS
VvSCP25 HAALESV-AGKWELCTDRI---------LY---HHDAGSMIKYHKNLTSN-GY-RALIFS
SCPLe HAEPATV-IGPWELCTDKI---------DL---DHDSGSMIPYHKNLTAR-GY-RAIIFS
AsSCPL1 GVREGTIGDGNWALCPEVPK-------LHL---TNDVPTTVPYHRRLTQR-GY-RALVYN
DgSCPL1 KIRKGMV--QEWVRCNKEDL--------HY---VYDVKASTGYHLNLSRK-GY-RSLIYS
DgSCPL2 HIRKDTV--PEWIRCNQEGL--------NY---LYDNISSIGYHLYLSRK-GC-RSLIYS
VvSCP31 HIRKNTI--REWQRCAM-GL--------SY---TPEIESSFEYHVTLSKK-GY-RSLIYS
CtAT1 HIRKGSI--GKWRRCTR-NL--------PY---TEDIPSSFEYHVNLSGK-GYYRSLVYS
AtSCT GVKK--V--GKWNRCNSQNI--------PY---TFEIFNAVPYHVNNSLK-GF-RSLIYS
BnSCT1 GVKKG-F--GKWSRCNTQNI--------PY---TYDIHNAIPYHVNNSRK-GF-RALIYS
BnSCT2 GVKKG-F--GKWSRCNTQNI--------PY---TYDIHNAIPYHVNNSRK-GF-RALIYS
AtSCPL17 LVNEGTV--RKWIRCNT-EI--------AY---NKDIRSSVPYHKYISIE-GY-RSLVFS
AtSMT HIEKGSK--GKWARCNR-TI--------PY---NHDIVSSIPYHMNNSIS-GY-RSLIYS
AtSAT HVTKGTK--GQWQRCNW-TI--------PY---DNNIISSVPYHMDNSIN-GY-RSLIYS
AtSST HVDKGSI--GEWIRDHR-GI--------PY---KSDIRSSIPYHMNNSIN-GY-RSLIFS
VvSCP27 YIRPGTV--EEFFRCNISL---------SY---TENVNNVFGYHKNLTNS-GL-RVLVFS
VvSCP45 YIRPGTV--EEFFRCNSSL---------SY---TENVNNVFGYHKNLTNS-GL-RVLVFS
GAT1 NVRVGTV--KYWSRCNKSL---------SY---TKDVQSVIDVHRYLSKK-QL-EVLVEV
DkSCPL2 HVRQGTV--AYWMRCNFSL---------SY---TKDIHSVVSVHEYLKTI-AL-QVLVAS
CsSCPL HVRKGSV--LNWERCNKSL---------SY---TKDILTVVPVHEELKEL-GL-EVLVET
VvSCP53 NVRNGTK--GNWQPCNSSLT--------GY---TEDVTTTLAYHRNFSHTSSL-RALIYS
VvSCP56 HVREGTK--GYWKRCNISGL--------AY---TEDVISSVAYHRNLSKT-GL-RALIYS
VvSCP52 RVREGTK--GHWVRCNITNL--------AF---TKDVTSTVAYHQNLTNT-GL-RALIYS
VvSCP57 HVREGTK--GHWVRCNVSGL--------AY---TRDVKSSIPYQRNLTQT-GL-RALIYS
SlGAC NVREGTT--LEWVRCNESMHYRGKERTESY---VYDVPSVIDDHQHLTSK-SC-RALIYS
GAT2 HVQPGTV--KTWRRCPKSFP--------SY---TENVDSTVAYHKNFTRT-GL-RALIYS
DkSCPL1 GVKEGTM-NSTWRRCAKSLP--------FY---EEDVSSTVAYHKNFTRT-AL-RALIYS
VvSCP60 HVRNGTI--PFWKRCNKTL---------DY---DSNVVSTVPYHRNLSDL-GY-RALIYS
VvSCP5 GIREGTI--PEWVRCNNSL---------AY---THDVFSTVAYIQKLHEK-GY-GGLIYS
VvSCP61 HVRE-AI--KDWVRCNESL---------SY---TSNVFSSVDYHRNLTKK-AY-RALIYS
VvSCP26 HANTTH-LPYYWDFCSGPLV-------YQV---DNLDMDLLPLIAYLLEQ-NI-RILLYS
VvSCP58 HANTTH-LPYHWGFCAGPLE-------YQI---DNLDMNLIPLIEHLIKE-GI-PILLFS
VvSCP10 HANRTG-LNYRWTMCSGVLN-------YSE---TDGNIDILPLLKRIVQN-SI-PVWVFS
VvSCP42 HANRTN-LPYHWTTCSNILF-------YNE---GDSNLDMLPLLKRILQD-KI-PVWIFS
VvSCP13 HAHLKG-VS-RWSICSEVLK-------YEY---RNLEIPTIHVVGAVLKS-GI-RVLVYS
VvSCP24 HARLVG-IR-SWTVCSDILD-------YEL---LNLEIPTISIVGSLIKA-GI-PVLVYS
VvSCP55 HARLVG-VN-KWSVCSNILD-------YEL---LDLEIPTISIVGKLIKA-GI-PVLVYS
VvSCP3 HAQLVG-TP-NWTLCYPDSA-------HFL---RDAVIPSINVVEWLVWS-GI-RASVYS
VvSCP4 HAQLVG-TP-NWTLCYPDSA-------HFL---KDAVIPSINVVEWLVRS-GI-RASVYS
VvSCP48 HARLVG-VA-NWSMCSGALR-------YNI---KDKEITMIPVMGSLVKS-GI-RTFVYS
VvSCP37 HARLVG-VA-NWSMCSGALR-------YNI---KDKEITMIPVMGSLVKS-GI-RTFVYS
VvSCP59 HARLVG-VA-NWSMCSGALR-------YNI---KDKEITMIPVMGSLVKS-GI-RTFVYS
VvSCP51 HANRTK-IPYAWEVCSS-VI-------TSW---TDSPSTMFPIYKRLISS-GL-QILIYS
VvSCP40 HANVTK-LKYDWEICNN-VV-------YNW---TDSAWSIITLLHEFMEN-GL-RVWVYS
VvSCP8 HANVTK-LKYEWRPCSD-ID-------KNW---TDSPLTIIPLLREFMAN-GL-RVWVFS
VvSCP20 HANVTK-LKYDWEPCSD-VI-------QNW---TDSPSTIIPLLHEFMEN-GL-RVWVFS
VvSCP33 HANVTN-IPGPWESCNDAIF-------YGW---KDMPLTVLPVIEELMVS-GI-SVWIYS
VvSCP29 HARN-----TTWGACSG-V---------GW---TDSPTTILPTIKQLMAS-GI-SVWIYS
MtSCP1 HAKP-----TNWTHCTH-LL-------TTW---KDSPATVLPTVKYLIDS-GI-KLWIYS
VvSCP36 HAATH----TKWEVCSDSVF-------HAY---HYTVFSVLPIYTKLIKA-GL-RIWVYS
VvSCP16 HANVTN-IPYPWTHCSNNI--------SFW---NDAPASILPIIKKLVDG-GL-RIWVFS
VvSCP21 HANVTK-LPYPYTTCSKVI--------RRW---NDSPDTVLPTIQKLLKA-GL-RIWVYS
VvSCP44 HANVTK-LPYPYTPCSNVI--------RKW---NDSAETMLPTIQKLLKA-GL-RIWVYC
VvSCP54 HVSDGH-RVKNWSICNADIF-------GNW---SQSQPSVLPIYRKLIAG-GL-RIWVYS
SbHNL1 HANVSGIVEYPWTVCSNTIF-------DQW---GQAADDLLPVYRELIQA-GL-RVWVYS
VvSCP50 HANTTK-IPYGWTACSEVLN-------RNW---NDTAESVLPIYREMIAA-GL-RVWVFS
VvSCP35 HANSTG-IPYKWTACSGVLI-------KYW---NDSEASMLPIYKELIEA-GL-RIWVFS
BRS1 HANVTG-IRYKWTACSDVLI-------KTW---KDSDKTMLPIYKELAAS-GL-RIWIFS
HvCPDWII HVNPVI-GKSKWETCSEVIN-------TNW---KDCERSVLHIYHELIQY-GL-RIWMFS
VvSCP14 HVYVDN-ATFKWATCSDEVS-------TTW---KDSPRSVLNIYRELIHA-RL-RIWIFS
VvSCP15 HVYTDN-APSKWATCSDEVS-------ATW---KDSPKTVLDVYRELIHA-GL-RIWIFS
TaCPDWII HANVTGAMNYTWATCSDTIN-------THW---HDAPRSMLPIYRELIAA-GL-RIWVFS
VvSCP38 HANVTH-LPYSWATCSSIVR-------RNW---SDSPKSMLPIFKQLISA-GI-RIWLFS
VvSCP39 HANVTQ-VSYPWRTCSNIVG-------IYW---ADSPLSMLPIYQELIAA-GL-RIWVFS
DgSCPL3 HANVTG-IPYPWATCNDIIG-------DNW---KDSPRSMLPIYRELIAA-GI-RIWVFS
VvSCP63 IIPKNV----SWGGQSDLVF-------ST--LSGDFMKPRIKEVDELLAK-GV-NVTIYN
VvSCP7 GANVSI----AWEDCSDVVG-------EA--LHEDVMKSVKFMV-ELLVK-KS-KVLLYQ
VvSCP19 GANVSI----AWEDCSDVVG-------EA--LHEDVMKSVKFMV-ELLVK-KS-KVLLYQ
ScCPY GAEV-D----HYESCNFDIN-------RNFLFAGDWMKPYHTAVTDLLNQ-DL-PILVYA
VvSCP18 GVGD-V----EFVSCSTVVY-------DA--MQRDWMRDMEVGIPALLED-GI-KMLIYA
VvSCP34 GVGD-M----EFVSCSSEVY-------NA--MQRDWMRDMEVGIPALLED-GI-KMLIYA
HvSCPIII GVGD-I----EFVSCSTSVY-------QA--MLTDWMRNLEVGIPALLED-GI-NVLIYA
VvSCP46 GVGD-I----EFVPCSPAVY-------QA--LLMDWMRNLEVGIPDLLED-GI-KLLVYA
VvSCP47 GVGD-I----DFVSCSPTVY-------QA--MLMDWMRNLEVGIPALLED-EV-KLLVYA

```
VvSCP32       GDHDMCVPFTGSEAWTR----SLGYKIVDE-----WRAWISN---DQVAGYTQGYEHG--
VvSCP62       GDHDMCVPFTGSEAWTR----SLGYKIMDE-----WRAWISN---DQVAGYTQGYEHG--
HvSCPI        GDHDMCVPFTGSEAWTK----SLGYGVVDS-----WRPWITN---GQVSGYTEGYEHG--
VvSCP25       GDHDMCVPYTGSQAWTR----SVGYKVVDE-----WRPWFFD---EQVAGYVQGYENN--
SCPLe         GDHDMCVPFTGSAVWTK----SLGYPIVDE-----WRPWYVN---DQVAGFIQGYANN--
AsSCPL1       GDHDLLMTHIGTHAWIR----SLGYPVVAP-----WRAWYSN---NEVAGFTVEYSNN--
DgSCPL1       GDHDMIVPYMGTEAWIK----SLNYSIIDD-----WRSWFVG---GQVAGYTRTYANN--
DgSCPL2       GDHDFLVPFLSTETWIT----SLNYSITDD-----WRPWFVG---GQVAGYTRSYANN--
VvSCP31       GDHDMIVPFFSTQAWIR----SLNYSIVDD-----WRSWMVE---GQVGGYTRTYSNQ--
CtAT1         GDHDLMVPFLGTQAWIR----SLNYSIVDD-----WRQWNTN---GQVAGYTRTYSNR--
AtSCT         GDHDSMVPFSSTQAWIR----ALNYSIVDD-----WRPWMMS--SNQVAGYTRTYANK--
BnSCT1        GDHDMMIPFSSTEAWIK----SLNYSIVDD-----WRPWMMN--SNQVAGYTRTYANK--
BnSCT2        GDHDMMIPFSSTEAWIK----SLNYSIVDD-----WRPWMMT--SNQVAGYTRTYANK--
AtSCPL17      GDHDMLVPFLGTQAWIR----SLNYSIVDD-----WRPWMVQ---NQVAGYTRTYANK--
AtSMT         GDHDIAVPFLATQAWIR----SLNYSPIHN-----WRPWMIN---NQIAGYTRAYSNK--
AtSAT         GDHDITMPFQATQAWIK----SLNYSIVDD-----WRPWMIN---DQIAGYTRTYSNK--
AtSST         GDHDITMPFQATQAWIK----SLNYSIIDD-----WRPWMIK---GQIAGYTRTYSNK--
VvSCP27       GDHDMVIPHVGIEQWIK----SLNISLGSD-----WRPWFVD---GQIGGYTRKYVNK--
VvSCP45       GDHDMVIPHVGIEQWIK----SLNISLGSD-----WRPWFVD---GQIGGYTRKYVNK--
GAT1          GDRDLVVPYPGAVEWIR----LLNLTIVSP-----WRPWFVD---GEIAGYTEKHSQN--
DkSCPL2       GDRDMVVPFVGTVKWIK----ALDLSVSEY-----WRPWFLD---GQVQGYTEKYDNN--
CsSCPL        GDRDMVVPFVGTVKWIK----SLNLTVVND-----WRPWFVD---GQVAGYTINYSEH--
VvSCP53       GDHDMSIPNIGTQEWIR----SLNMTLADT-----WRGWMVD---AQVAGYTKRYTYG--
VvSCP56       GDHDMSVPHIGTQQWID----SLNLTLADT-----WRAWYTE---GQVAGYTKRYTND--
VvSCP52       GDHDMSIPHIGTQEWIN----SLNLTLEDP-----WRTWCTD---GQVAGYTETFTND--
VvSCP57       GDHDMSIPHVGTQEWIN----LLNLTLADT-----WRAWYTD---AQVSGYTQRFTND--
SlGAC         GDHDMVVPHLSTEEWIE----TLKLPIADD-----WEPWFVD---DQVAGYKVKYLQN--
GAT2          GDHDLSTPYIGTLEWIK----SLDVPVFDK-----WRPWYVD---GQIAGYTTKFMND--
DkSCPL1       GDQALSIPYLGTLEWIN----SLGVPIFDT-----WRPWFVD---GQVAGYTQKYEKN--
VvSCP60       GDHDMLIPYVGTERWVK----SLNISVLNG-----WEPWFVD---GQVAGYSVVYQAN--
VvSCP5        GDHDMLVPHMGTQEWIN----SLNLSISKD-----WEPWFVD---GQVAGFSIEYSNS--
VvSCP61       GDHDMLIPYVGTQAWIA----SLNLNISED-----WQPWFVD---GQVAGFRVEYLHN--
VvSCP26       GDQDAKVPLTQTRLITNNLAKDLKLVPFTK-----YGTWYDK---EQVGGWSQSFGRLRD
VvSCP58       GDQDAIIPLTQTRIIANNVAKDLKLVPFTE-----YGTWYDK---KQVGGWTQSFGGLRE
VvSCP10       GDQDSVVPLLGSRTLIRELAQEMKFKITVP-----FGAWFHK---GQVGGWATEYGN---
VvSCP42       GDQDSVVPLMGSRTLVRELAKDLNFQHTVP-----YGAWFHK---GQVGGWQTEYGN---
VvSCP13       GDQDSVVPLTGTRTLVNGLAKDLGLNTTVP-----YRNWFQG---RQVGGWTQVYGD---
VvSCP24       GDQDSVIPLTGSRTLVHNLAKELGLNTTVP-----YRVWFEG---KQVGGWTRVYGN---
VvSCP55       GDQDSVIPLTGSRTLVHGLAEELGLNTTVP-----YRVWFEG---KQVGGWTQVYGN---
VvSCP3        GDQDSRISLIGTRSLLEGLAKKLKLKTTVP-----YRNWFEK---KQVGGWTQVYGD---
VvSCP4        GDQDSRMSLFGTRSLLEGLAKKLKLKTTVP-----YRNWFEK---KQVGGWTQVYGD---
VvSCP48       GDQDSVIPLFGTRTLVDGLAKELRLNTTVP-----YRNWFEG---EQVGGWTQVYGD---
VvSCP37       GDQDSVIPLFGTRTLVDGLAKKLRLNTTVP-----YRNWFEG---EQVGGWTQVYGD---
VvSCP59       GDQDSVIPLFGTRTLVDGLAKELRLNTTVP-----YRNWFEG---EQVGGWTQVYGD---
VvSCP51       GDVDAVVSVVGTRYSIN----ALNLKVIRP-----WHPWSES--TKVVGGYRVVYEG---
VvSCP40       GDVDGRVPVTSTLASLA----KMRLTVKTP-----WHPWFLH---GEVGGYTEVYKG---
VvSCP8        GDTDGDVPVTSTMASIG----KMRLSVKTP-----WHPWFVA---GEVGGYTEVYKG---
VvSCP20       GDTDGRVPVTSTMASID----TMKLSVKTP-----WHPWFVA---GEVGGYTEVYKG---
VvSCP33       GDTDGRVPTTSTRYSIN----NLGTSVKTP-----WYPWYTQ---GEVGGYAVGYKN---
VvSCP29       GDTDGRVPVTSSRYSIN----TFKLPVKTA-----WRPWYYN---KEVGGYVVEYKG---
MtSCP1        GDTDVVVPTTSSRYLIN----TLKLPINSA-----WRPWYSG---KEIGGYVVGYKG---
VvSCP36       GDTDGRVPAIGTRYCVE----ALGLPLKAP-----WRSWYHH---HQVGGRIVEYE----
VvSCP16       GDTDGRIPVSSTRLTLR----KLGLKTIQE-----WTPWYTS---HEVGGWTIEYD----
VvSCP21       GDTDGRVPVTSTRYSIN----KMGLRIQQK-----WRAWFDR---KQVAGWVVTYEG---
VvSCP44       GDTDGRVPVTSTRYSIN----KMGLRIQKG-----WRAWFHR---KQVAGWVVTYEG---
VvSCP54       GDTDGRVPVLSTRYCLS----TLKLPITRA-----WRPWYHQ---QQVSGWFQEYK----
SbHNL1        GDTDSVVPVSSTRRSLA----ALELPVKTS-----WYPWYMAPTEREVGGWSVQYE----
VvSCP50       GDVDSVVPVTATRYSLA----HLKLATKIP-----WYPWYVK--K-QVGGWTEVYE----
VvSCP35       GDTDAVVPVTATRFSLN----HLNLTVKTP-----WYPWYSG--G-QVGGWTEVYE----
BRS1          GDTDSVVPVTATRFSLS----HLNLPVKTR-----WYPWYTD--N-QVGGWTEVYK----
HvCPDWII      GDTDAVIPVTSTRYSID----ALKLPTVTP-----WHAWYDD--DGEVGGWTQGYK----
VvSCP14       GDTDAVIPVTSTRYSID----ALKLPTVSP-----WRAWYD---DGQVGGWTQDYA----
VvSCP15       GDTDAIIPVTSTRYSVD----ALKLPTVGP-----WRAWYD---DGQVGGWSQEYA----
TaCPDWII      GDTDAVVPLTATRYSIG----ALGLPTTTS-----WYPWYDD--Q-EVGGWSQVYK----
VvSCP38       GDTDAVLPLTATRYSIK----ALKLKTITN-----WHAWYDD--KQEVGGWSQVYE----
VvSCP39       GDTDAVVPVTATRYSID----ALKLPTITN-----WYAWYDN--H-KVGGWSQVYK----
DgSCPL3       GDTDSVVPVTATRYSLD----TLELPTLVK-----WHPWYDN--G-QMGGWSEVYD----
VvSCP63       GQLDLICATMGTEAWVE----KLKWDGLKEFLSMKRTPLYCG-GEGGTKGFTKSYKN---
VvSCP7        GQFDLRDGVVSTEAWMK----SMKWEGIDKFQAAERKVW--E-VKGELAGYVQKWGS---
VvSCP19       GQFDLRDGVVSTEAWLK----TMKWEGIDKFQAAERKVW--K-VKGELAGYVQKWGN---
ScCPY         GDKDFICNWLGNKAWTD----VLPWKYDEEFASQKVRNWTAS-ITDEVAGEVKSYKH---
VvSCP18       GEYDLICNWLGNSNWVH----AMKWSGQKDFEASPTVP--YL-VDGKEAGQLKNHGP---
VvSCP34       GEYDLICNWLGNSNWVH----AMKWSGQKDFEASPTVP--YL-VDGKEAGQLKYHGR---
HvSCPIII      GEYDLICNWLGNSRWVH----SMEWSGQKDFAKTAESS--FL-VDDAQAGVLKSHGA---
VvSCP46       GEYDLICNWLGNSRWVH----AMEWGGQLEFQAAPEVP--FV-IGDSKAGLMKIHGP---
VvSCP47       GEYDLICNWLGNSRWVH----AMKWSGQKDFQASLEIP--FE-VRDSHAGLVKSYGP---
```

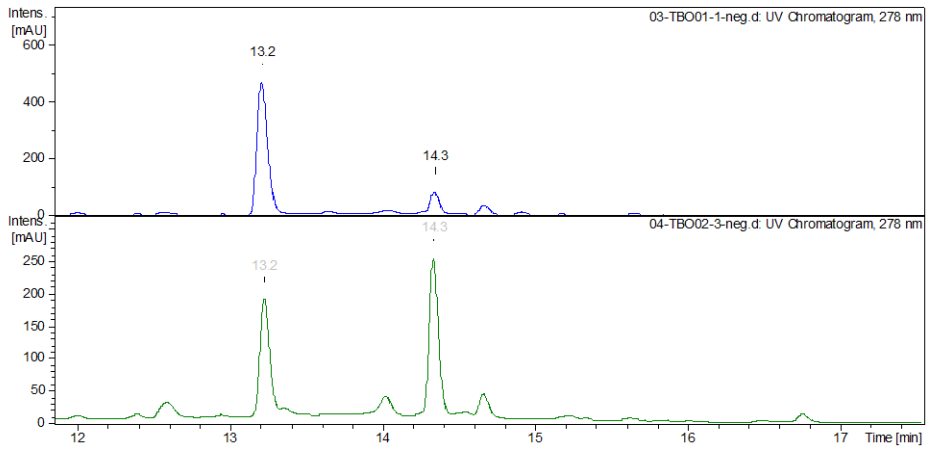

VvSCP32 ------LTFLTIKGAGHTVPEYKPKEALDFF--SRWLDGKAI*----------
VvSCP62 ------LTFLTIKGAGHTVPEYKPREALDFF--GRWLEGKAI*----------
HvSCPI ------LTFATIKGAGHTVPEYKPQEAFAFY--SRWLAGSKL----------
VvSCP25 ------LTFLTVKGSGHTVPEYKPREALAFY--SRWLTGRPI*----------
SCPLe ------LIFMTIKGAGHTVPEYKPREALAFY--SRWLEGKKI----------
AsSCPL1 ------LTFATVKGAGHMAPESRPKQCLDMV--RRWISPAGKL----------
DgSCPL1 ------MTFATVKGGGHTAPEYKPCECLAML--KRWLSYESL----------
DgSCPL2 ------MTFATVKGAGHTAPEYKPYECFSMM--KKWLSYEPL----------
VvSCP31 ------MTFATVKGGGHTAPEYRPKECFGMY--KRWVSGQPL*----------
CtAT1 ------MTFATVKGGGHTAPEFKPEECFAMY--SRWISKRPL----------
AtSCT ------MTFATIKGGGHTA-EYTPDQCSLMF--RRWIDGEPL----------
BnSCT1 ------MTFATIKGGGHTA-EYNPDQCSLMF--KRWIDGESL----------
BnSCT2 ------MTFATIKGGGHTA-EYNPDQCSLMF--KRWIDGESL----------
AtSCPL17 ------MTFATVKGGGHTS-EYKPVETYIMI--KRWLSGQPL----------
AtSMT ------MTFATIKGGGHTA-EYRPNETFIMF--QRWISGQPL----------
AtSAT ------MTFATVKGGGHTA-EYLPNESSIMF--QRWISGQPL----------
AtSST ------MTFATVKGGGHTA-EYLPEESSIMF--QRWISGQPL----------
VvSCP27 -A--YSLTYSTLKGAGHSPTEYRRRESYEMF--YRWIHHYPL*----------
VvSCP45 -A--YSLTYSTLKGAGHSPTEYRRRESYEMF--YRWIHHYPL*----------
GAT1 -G--YRLTYATVKGAGHTAPEYYRRECYYMF--KRWVHYYPLQ----------
DkSCPL2 -G--YYLTYVTVKGAGHTAPEYHRKQVYHLF--DRWIHYYPV----------
CsSCPL -G--YRLTYATVKGAGHTAPEYYRMECYYMF--DRWTHYYPI----------
VvSCP53 -D--FSLTYATVKGAGHIPATYKTRQCYEMI--ERWLAHYPLQ----------
VvSCP56 -D--FALTYATVKGAGHVAPEYKPQQCYAML--KRWFAYYQL*----------
VvSCP52 -D--FDLTFATVKGAGHVAIEYKPKECYAMI--DRWFAHYPL*----------
VvSCP57 -D--FSLTFATVKGAGHVAIEYKAKECYAMI--KRWFGYYPL*----------
SlGAC -D--YEMTYATVKGAGHTAPEYKPEQCLPMV--DRWFSGDPL----------
GAT2 -H--YRLTYATLKGAGLTAPEYKHKESLALV--DRWFAYYPIQ----------
DkSCPL1 -S--YSLTFATVKGAGETAPEYKRKEALAMV--NRWFAGYPV----------
VvSCP60 -KTESDITYATVKGGGHTAPEFRPKQCLAMI--DRWLAFYPL*----------
VvSCP5 -K--RGMTFATVKGGGHTAPEYKPKECLAMI--YRWLAYYPL*----------
VvSCP61 -K--YLMTFATVKGAGHTAPEYRPKEGFAMV--YRWLAHYFL*----------
VvSCP26 GMNLTLLTFATVRGAAHEVPFTSPSQALTLF--KSFLSGSPPPRLHE*----------
VvSCP58 GKNVTLLTFATVRGAAHEVPFTSPSQALTMF--KSFLSGSPLPRPHE*----------
VvSCP10 -----LLTFATVRGAAHMVPYAQPSRALHLF--SSFVRGRRLPNTTRPSIDD-Q------
VvSCP42 -----LLTFATVRGAAHMVSYSQPSRALHLF--ATFIHGRRLPNNTRPSID*---------
VvSCP13 -----KLSFATIRGASHEAPFSQPERSLVLF--NTFLQGKPLPEATVVL*----------
VvSCP24 -----ILSFATIRGASHEAPFSQPERSLVLF--KSFLEARPLPEFF*----------
VvSCP55 -----ILSFATIRGASHEAPFSQPERSLVLF--RAFLGGRPLPQAF*----------
VvSCP3 -----ILSFATIRGGSHTAPISQPTRSLALF--TAFLEGKPLPDA*----------
VvSCP4 -----ILSFATIRGGSHTAPISQPARSLALF--TAFLEGKPLPDA*----------
VvSCP48 -----ILSFATVRGGSHTVPGTQPARVSVGAHFRVVLRSAQVSELAKEASVW-PRWLISD
VvSCP37 -----ILSFATVRGGSHTVPGTQPARALVLFT--AFLKGQPPPAE*----------
VvSCP59 -----ILSFATVRGGSHTVPGTQPARALVLFT--AFLKGQPPPAE*----------
VvSCP51 ------LTFATIRGAGHEVPRFQPRRAFALM--ESFVAGK*----------
VvSCP40 -----DLTFATVRGAGHQVPSFQPRRALSFI--IHFLAGTPLPN*----------
VvSCP8 -----DLTFATVRGAGHQVPSFRPKRALSLI--VHFLSGTPLPKPSSKH*----------
VvSCP20 -----DLTFATVRGAGHQVPSFRPKRALSLI--SHFLSGTPLPRRSSKH*----------
VvSCP33 ------LSFVTIRGAGHFVPSYQPTRALAFF--SSFLAGKLPSADEP*----------
VvSCP29 ------VVFATVRGAGHLVPSYQPGRALTMI--ASFLQGTLPPPS*----------
MtSCP1 ------LTFVTVRGAGHLVPSWQPERALTLI--SSFLYGILPSGSPSN----------
VvSCP36 -----GLTYLTVRGAGHLVPLNKPSQAFALI--HSFLTAIQLPTRK*----------
VvSCP16 -----GLTFVTVRGAGHEVPTFAPKQAFQLI--RHFLDNEKLPSTPIQ----------
VvSCP21 -----GLTLATVRGAGHQVPILAPSQSLALF--SHFLSDATLPSSRAQ----------
VvSCP44 -----GLTLATVRGAGHQVPILAPAQSLALF--SHFLSAANLPSSRA*----------
VvSCP54 -----GLTFATFRGAGHAVPVFKPSESLAFF--SAFLQGESPPCQS*----------
SbHNL1 -----GLTYVSPSGAGHLVPVHRPAQAFLLF--KQFLKGEPMPAEEKNDILL-PSEKAPF
VvSCP50 -----GLTFATVRGAGHEVPLFKPRAALELF--KSFLRGLPLPKS*----------
VvSCP35 -----GLTFATVRGAGHEVPLFQPMRAFLLF--RSFLGGKQLPSSSEN*----------
BRS1 -----GLTFATVRGAGHEVPLFEPKRALILF--RSFLAGKELPRSY----------
HvCPDWII -----GLNFVTVRGAGHEVPLHRPKQALTLI--KSFLAGRPMPVLSDLRSDM--------
VvSCP14 -----GLTFVTVRGAGHEVPLHKPKQAFTLF--KAFLSGAPMPYMEQVSYS*--------
VvSCP15 -----GLTFVTVRGAGHEVPLHKPKQALTLI--NAFLKGTSMPSLEQLADS*--------
TaCPDWII -----GLTLVSVRGAGHEVPLHRPRQALVLF--QYFLQGKPMPGQTKNAT----------
VvSCP38 -----GLTFTTVRGAGHEVPLGQPRRALILL--GHFLNNKPMPAAPTLF*----------
VvSCP39 -----GLTFVTVTGAGHEVPLHRPRQAYILF--RSFLENKPMPS*----------
DgSCPL3 -----GLSLIVVRGAGHEVPLHAPRKAYIMF--KSFLENKPMPH----------
VvSCP63 ------LHFYWILGAGHFVPVDQPCIALNMV--GGITHSPMASIS*----------
VvSCP7 ------LSHVVVSGAGHLVPADQPVNSQIMI--EDWVLERGLFGNEQERN*----------
VvSCP19 ------LSHVVVSGAGHLVPADQSVNSQIMI--EDWVLERGLFANEQEANPSSSSTGNGF
ScCPY ------FTYLRVFNGGHMVPFDVPENALSMV--NEWIHGGFSL----------
VvSCP18 ------LAFLKVHNAGHMVPMDQPKAALQML--KTWTQGKLAPIETKDD*----------
VvSCP34 ------LAFLKVHNAGHMVPMDQPKAALQML--KTWTQGKLAPKIKDETT-----W*---
HvSCPIII ------LSFLKVHNAGHMVPMDQPKAALEML--RRFTQGKLKEAVPEEES-----STTSF
VvSCP46 ------LTFLKVHDAGHMVPMDQPRVALEML--KRWFENKLPENTPAESK-----EPEKR
VvSCP47 ------LTFLKVHDAGHMVPMDQPEASLEML--KRWMEGKLVEGQ-DESE-----EPEKL

```
VvSCP32       --------------------
VvSCP62       --------------------
HvSCPI        --------------------
VvSCP25       --------------------
SCPLe         --------------------
AsSCPL1       --------------------
DgSCPL1       --------------------
DgSCPL2       --------------------
VvSCP31       --------------------
CtAT1         --------------------
AtSCT         --------------------
BnSCT1        --------------------
BnSCT2        --------------------
AtSCPL17      --------------------
AtSMT         --------------------
AtSAT         --------------------
AtSST         --------------------
VvSCP27       --------------------
VvSCP45       --------------------
GAT1          --------------------
DkSCPL2       --------------------
CsSCPL        --------------------
VvSCP53       --------------------
VvSCP56       --------------------
VvSCP52       --------------------
VvSCP57       --------------------
SlGAC         --------------------
GAT2          --------------------
DkSCPL1       --------------------
VvSCP60       --------------------
VvSCP5        --------------------
VvSCP61       --------------------
VvSCP26       --------------------
VvSCP58       --------------------
VvSCP10       --------------------
VvSCP42       --------------------
VvSCP13       --------------------
VvSCP24       --------------------
VvSCP55       --------------------
VvSCP3        --------------------
VvSCP4        --------------------
VvSCP48       WGQGNKHFDGTWRSTYKEE*
VvSCP37       --------------------
VvSCP59       --------------------
VvSCP51       --------------------
VvSCP40       --------------------
VvSCP8        --------------------
VvSCP20       --------------------
VvSCP33       --------------------
VvSCP29       --------------------
MtSCP1        --------------------
VvSCP36       --------------------
VvSCP16       --------------------
VvSCP21       --------------------
VvSCP44       --------------------
VvSCP54       --------------------
SbHNL1        Y-------------------
VvSCP50       --------------------
VvSCP35       --------------------
BRS1          --------------------
HvCPDWII      --------------------
VvSCP14       --------------------
VvSCP15       --------------------
TaCPDWII      --------------------
VvSCP38       --------------------
VvSCP39       --------------------
DgSCPL3       --------------------
VvSCP63       --------------------
VvSCP7        --------------------
VvSCP19       YV*-----------------
ScCPY         --------------------
VvSCP18       --------------------
VvSCP34       --------------------
HvSCPIII      YAAM----------------
VvSCP46       VAQM*---------------
VvSCP47       VAQM*---------------
```

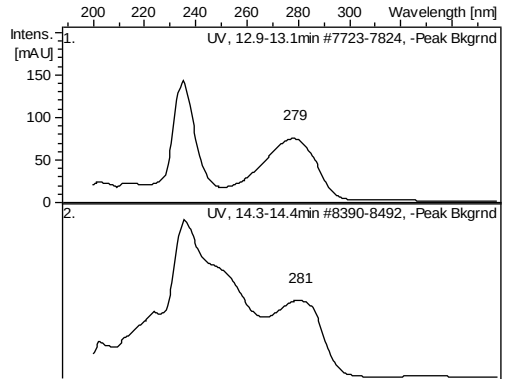
Supplemental Figure 4.Alignement of Serine carboxypeptidases protein sequences by ClustalW2.
Putative glycosylation sites are in white letters and highlighted in dark blue. The cysteins which constitute the disulfide bridges 1, 2 and 3 are highlighted in light blue, red and yellow, respectively.The catalytic residues are in yellow letters and highlighted in black. Putative His residue involved in VvGAT2 catalytic triad are in bold letters and highlited in yellow. The divergent amino acid in VvGAT1 between Maccabeu and Portan cultivars is in white letters and highlighted in dark green.
Grapevine sequences:
VvSCP3: VIT_08s0040g01110; VvSCP4: VIT_08s0040g01060; VvSCP5: VIT_03s0091g01240; VvSCP7: VIT_01s0011g04900; VvSCP8: VIT_13s0158g00070; VvSCP10: VIT_18s0001g15550; VvSCP13: VIT_08s0040g01140; VvSCP14: VIT_11s0037g00800; VvSCP15: VIT_11s0037g00740; VvSCP16: VIT_10s0042g01050; VvSCP18: VIT_06s0061g00520; VvSCP21: VIT_01s0011g02000; VvSCP24: VIT_17s0000g00370; VvSCP25: VIT_07s0141g00520; VvSCP26: VIT_14s0108g00110; VvSCP27: VIT_03s0088g00050; VvSCP29: VIT_18s0001g11100; VvSCP31: VIT_18s0001g08840; VvSCP32: VIT_04s0023g00700; VvSCP33: VIT_03s0063g00870; VvSCP34: VIT_06s0061g00530; VvSCP35: VIT_11s0037g00320; VvSCP36: VIT_05s0077g00560; VvSCP37: VIT_08s0040g01040; VvSCP38: VIT_13s0019g04040; VvSCP39: VIT_13s0019g04030; VvSCP40: VIT_13s0158g00050; VvSCP42: VIT_12s0055g00820; VvSCP44: VIT_01s0011g02030; VvSCP45: VIT_03s0088g00110; VvSCP46 and VvSCP47: VIT_13s0019g05130; VvSCP48: VIT_08s0040g01070; VvSCP50: VIT_14s0068g00260; VvSCP51: VIT_03s0063g00860; VvSCP52: VIT_11s0052g00770; VvSCP53: VIT_11s0065g00720; VvSCP54: VIT_10s0116g00740; VvSCP55: VIT_01s0244g00100; VvSCP56: VIT_11s0052g00790; VvSCP57: VIT_11s0052g00750; VvSCP58: VIT_14s0108g00100; VvSCP59: VIT_08s0040g01130; VvSCP60: VIT_03s0091g01200; VvSCP61: VIT_03s0091g01270.
Sequences from other organisms:
SlGAC: *Solanum lycopersicum* glucose acyltransferase (AAF64227); CtAT1: *Clitoria ternalea* Acyltransferase 1 (BAF99695, Patent Publication Number WO/2007/046148); AtSMT: *Arabidopsis thaliana* 1-*O*-sinapoyl-β-glucose:L-malatesinapoyltransferase (AAF78760); AtSCT: *Arabidopsis thaliana*1-*O*-sinapoyl-β-glucose:cholinesinapoyltransferase (AAK52316); BnSCT1: *Brassica napus*1-*O*-sinapoyl-β-glucose:cholinesinapoyltransferase 1 (AAQ91191); BnSCT2 (CAM91991); AtSAT: *Arabidopsis thaliana*1-*O*-sinapoyl-β -glucose:anthocyaninsinapoyltransferase (AEC07395); AtSST:*Arabidopsis thaliana* 1-*O*-sinapoyl-β -glucose:1-*O*-sinapoyl-β-glucose sinapoyltransferase (AEC07397); AsSCPL1: *Avena sativa* Saponin-deficient 7 (ACT21078); DkSCPL1: *Diospyros kaki* Serine Carboxypeptidase Like 1 (BAF56655); DkSCPL2 (BAH89272); AtSCPL17 (AAS99709); CsSCPL: *Camellia sinensis* SCPL (AIW39897). HvSCPI: *Hordeum vulgare* serine carboxypeptidase I (P07519); HvCPDWII: *Hordeum vulgare* serine carboxypeptidase II (P55748); HvSCPIII: *Hordeum vulgare* serine carboxypeptidase III (P21529); TaCPD-WII: *Triticum aestivum* serine carboxypeptidase II (P08819); ScCPY: *Saccharomyces cerevisae* carboxypeptidase Y (NP_014026); SbHNL1: *Sorghum bicolor* Hydroxynitrilelyase (CAD12888); BRS1: *Arabidopsis thaliana* brassinosteroid-insensitive 1(Q9M099); SCPLe: serine carboxypeptidase *Lycopersicum esculentum* (Q9M513); MtSCP1: *Medicago truncatula* serine carboxypeptidase 1 (XP_003592239)

Supplemental Figure 5.

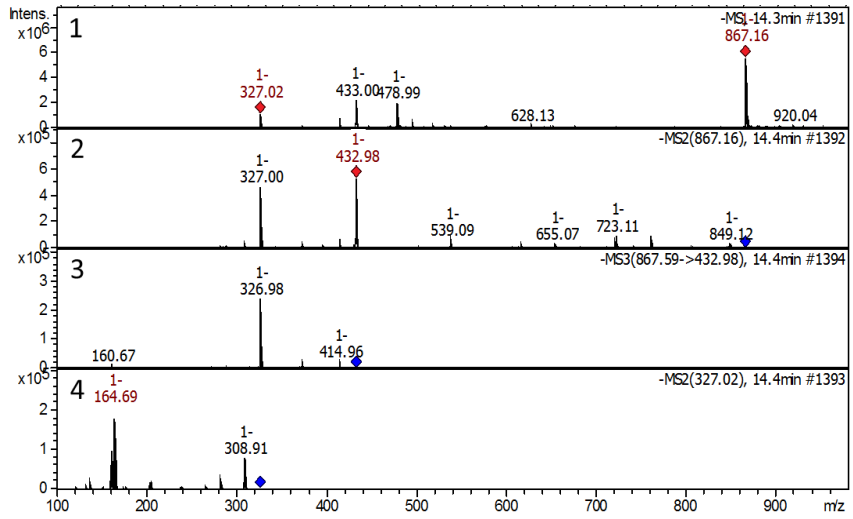
A



B



C



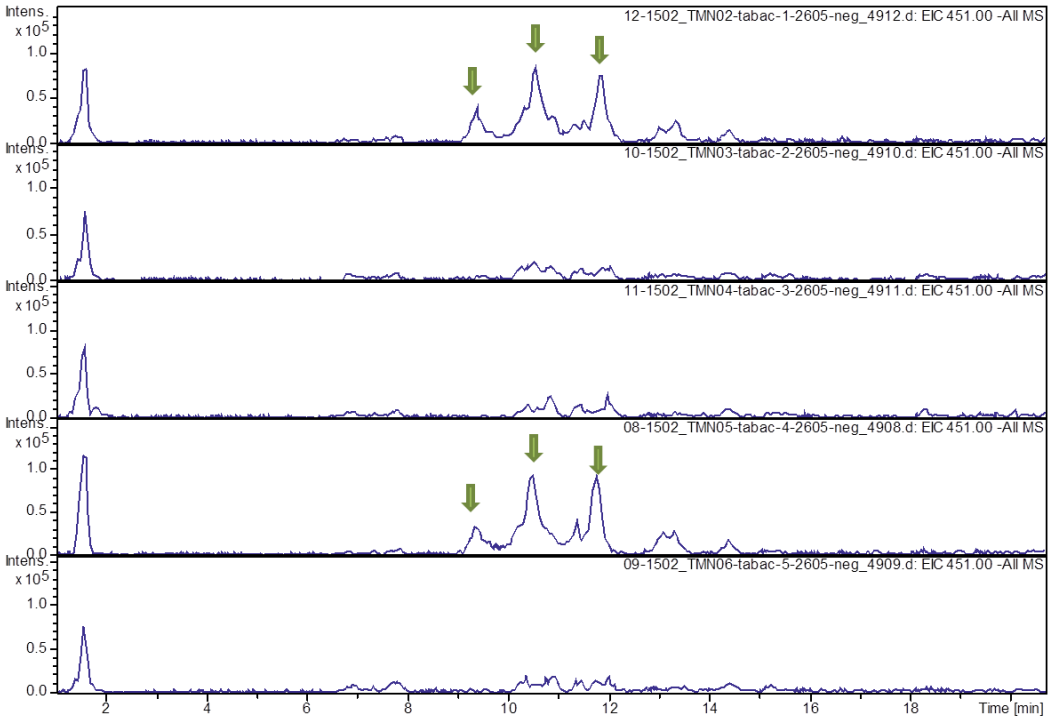
Supplemental Figure 5.UPLC-DAD-MS analysis of the unknown compound.

A. Chromatogram at 280 nm for enzymatic assay with proteins extracted from tobacco leaves inoculated with pEAQ-HT-GFP (control, blue line), pEAQ-HT-DEST1:GAT1 (green line). Peak at 13.2 min: epicatechin, peak at 14.3: unknown compound.

B. UV spectrum of epicatechin (upper chromatogram) and the unknown compound.

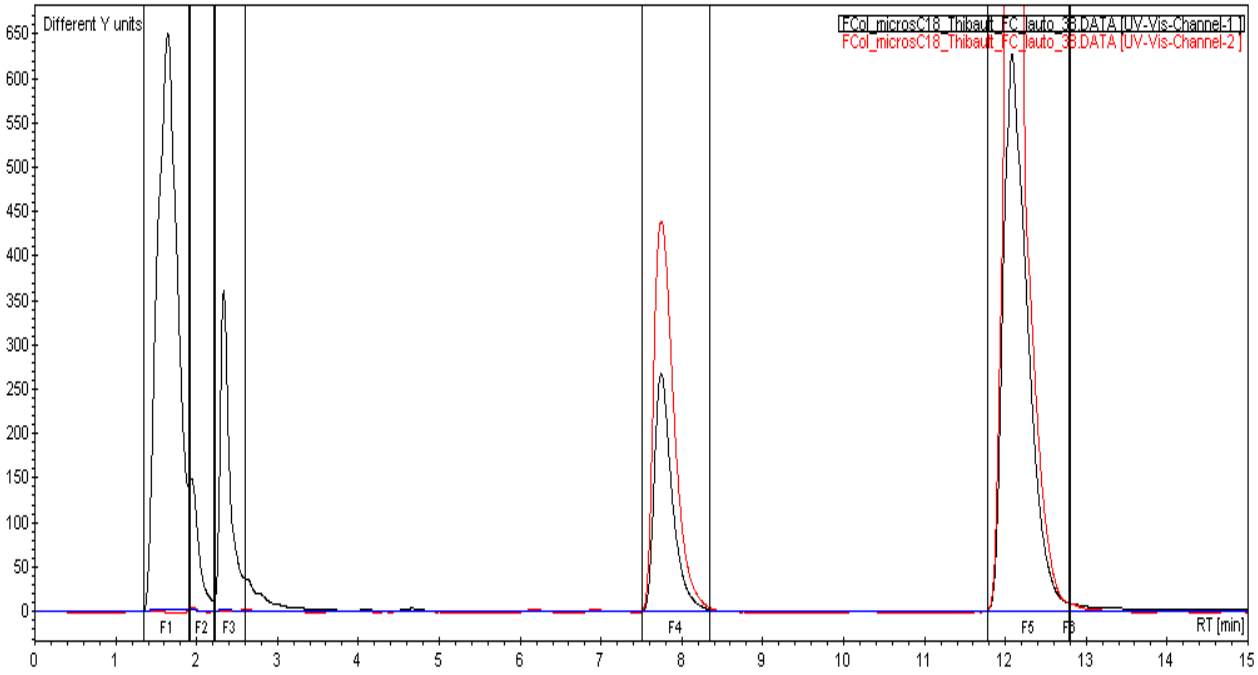
C. MS analysis in negative mode of the unknown compound. 1: full scan. 2: fragmentation spectrum of negative ion at 867. 3: fragmentation spectrum of negative ion at 432. 4: fragmentation spectrum of negative ion at 326.

Supplemental Figure 6.



Supplemental Figure 6. Extracted-ion chromatogram for epicatechin glucoside in tobacco leaves.
The arrows show the peaks that could correspond to epicatechin glucoside, detected at m/z 451 in the negative ion mode.
In conditions **A**, **B** and **D**, a mixture of (-)-epicatechin and β-glucogallin (final concentration 200 µM) has been infiltrated 3 days after inoculation in the same inoculated leaves zones. After substrates infiltration, the leaves have been detached and their petiole dipped into the mixture.
A. Non inoculated leaves infiltrated and dipped in the substrates mixture. **B**. Leaves inoculated with pEAQ-HT-GFP vector and infiltrated and dipped in the substrates mixture. **C**. Leaves inoculated with pEAQ-HT-GFP vector. **D**. Leaves inoculated with pEAQ-HT-DEST1:GAT1 construct and infiltrated and dipped in the substrates mixture. **E**. Leaves inoculated with pEAQ-HT-DEST1:GAT1 construct.

Supplemental Figure 7.



Supplemental Figure 7.Production of putative substrate for VvGATs.
Feruloyl glucose ester (Fraction 4, F4) been produced from ferulic acid (F5) and UDP-Glucose (F1) using *Vitis vinifera* Glucosyltransferase 2 (VvGT2) in the enzymatic assay conditions described in Khater *et al*., (2012). Semi-preparative HPLC conditions are described in Material and Methods.The black and the red line correspond to the chromatogram at 280 and 320nm, respectively. F4 has been collected.
RT: retention time (in minutes).

Supplemental Table 1.

Family / Species	Molecule	Concentration (nmol.g^{-1})	Organ/tissu	Ref
Amaranthaceae				
Beta vulgaris (beet)	GA	82.29 *	Leaves, red tissues	1
		70.54 *	Stem, red tissues	
		217.49 *	Leaves, white tissues	
		1.5 *	Stem, white tissues	
Anacardiaceae				
Mangifera indica L.	GA	10.23	Fruit/pulp	2
cv. 'Keitt' (mango)	β-G	116.89		
Mangifera indica L.	GA	5.29-17.63 *	Fruit/ pulp	3
Mangifera indica L.	GA	556.08-580.18 **	Fruit/ pulp	4
cv. Ataulfo				
Pistacia lentiscus	GA	21749 **	Leaves	5
Pistacia vera L.,	GA	74.42 *	Seed	6
var. Bronte		8542.85 *	Skin	
Asteraceae				
Cichorium sp.	GA	470-4570 *	?	7 †
Betulaceae				
Coryllus avelana (hazelnut)	GA	2.94-11.76	Fruit	8
Brassicaceae				
Brassica oleracea (cauliflower)	GA	39.97-41.15	?	9 †
Brassica oleraceae L.	GA	0.03 *	Mature leaves	10
var. *acephala*	GA	0.01 **	Mature seeds	
Brassica juncea	GA	2.94 **	Leaves	11
(potherb mustard)				
Clusiaceae				
Garcina mangostana	GA	1234.42	Fruit / pericarp	12
(mangosteen)				
Ebenaceae				
Diospyros kaki	GA	1123.32 **	Fruit/pulp	13
Euphorbiaceae				
Ricinus communis	GA	39971.78-67559.34	Leaves	14
(castor oil plant)				
Jatropha curcas	GA	88.94-443.16	Leaves	15
Fabaceae				
Glycine max L. Merr.	GA	4114.74 *	Seeds	17 †
(soybean)				
Phaseolus vulgaris	GA	103.3-125	Seeds (beans)	18
(black bean)				
Robinia pseudoacacia	GA	40.68-489.65	Heartwood	19
(acacia)				
Fagaceae				
Castanea sativa Mill.	GA	20338.58 - 57018.58 *	Fruit without skin	20 †
(chestnut)				
Quercus pyrenaica (oak)	GA	584.88 - 1058.08	Wood	21
Quercus robur	GA	1063.95 - 2880.32	Wood	22
Quercus petraea	GA	423.23 - 2645.19		
Quercus pyrenaica	GA	2221.03 - 2615.8		
Quercus faginea	GA	2562.9 - 3421.11		
Quercus alba	GA	564.31 - 605.45		
Iridaceae				
Crocus sativus (saffron)	GA	0.1	Dormant corn	23
Lamiaceae				
Origanum vulgare L. (oregano)	GA	605.45	Herb	24 †
Salvia officinalis L. (sage)	GA	617.21	Herb	
Malvaceae				
Gossypium hirsutum	GA	282.15 *	Leaves	25
(Mexican cotton)				
Theobroma cacao (cacao tree)	GA	34093.58	Shoot	26
Moraceae				
Ficus carica L. (fig)	GA	8.82-2.34 *	Whole fruits	27
Musaceae				

Musa sp. (Banana)	GA	33.51-76.42	Fruit	28 †
Myricaceae				
Myrica rubra (Bayberry)	GA	15.28-41.15 *	Berries	29
Myrtaceae				
Eucalyptus globulus	GA	20044.67-51904.54	Bark	30
Eucalyptus globulus	GA	30801.79	Fruit	31
Eucalyptus sp.	GA	35092.88-61897.48 **	Leaves	32
Syzygium aromaticum	GA	46055.73	Flower buds	34
(clove)		18222.43-35857.04	Flower buds	35 †
Oleaceae				
Olea europaea L. (olive)	GA	1.6-2.7	Fruit/pulp	36 †
Rosaceae				
Prunus domestica L. (plum)	GA	0.59	Seed	37
Rubus glaucus	GA	105.81	Berries	38
(andean raspberry)				
Rubus adenotrichus	GA	27.04 **	Berries	
Rubus laciniatus	GA	117.56	Berries	39 †
(evergreen blackberry)				
Rubus ursinus x *idaeus*		529.04		
Rubus ursinus		176.35		
Rubus chamaemorus L.	GA	246.88 *	Berries	40 †
Malus domestica (apple)	GA	124.03 *	Fruit / peel	41
		63.48	Fruit / flesh	
Malus domestica, Fuji	GA	55.84 **	Fruit / flesh	42
Prunus avium (wild cherry)	GA	7.17	Heartwood	43
Prunus persica	GA	9.72 *	Fruit / pulp	44
var. *platycarpa*				
Rutaceae				
Citrus grandis x *C. paradisi*	GA	0.36 *	Fruit/ pulp	45 †
(grapefruit)		0.08 *	Fruit/ skin	
Salicaceae				
Populus simonii x *pyramidalis*				
'Opera 8277'	GA	~587.82 **	Leaves	46
Sapotaceae				
Chrysophyllum cainito (cainito)	GA	9.41 *	Fruits	47
Solanaceae				
Solanum melongena L.	GA	7-11.87 *	Fruits	48 †
(aubergine)				
Solanum tuberosum (potato)	GA	88.17-511.4 *	Whole tuber	49
Solanum tuberosum	GA	3409.36	Tuber peel	50
var. Viking (red)				
Solanum tuberosum	GA	3703.27	Tuber peel	
var. Kennebec (brown)				
Solanum lycopersium,	GA	176.93**	Flesh	51
var. Jinpeng (tomato)				
Lycopersicon esculentum,	GA	199.98 **	Hairy root	52
var. Pusa early dwarf (tomato)				
Theaceae				
Camellia sinensis (tea plant)	GA	646.6	Buds	53
		117 - 352	Leaves	
		58.78	Stem	
	β-G	8378.33	Buds	
		651-1272	Leaves	
		124.12	Stem	

Supplemental Table 1. GA and β-glucogallin content in plants.

†: Reference found in Phenol Explorer, version 3.6.

Original data have been converted in nmol.g^{-1} using 170.12 and 322.26 g.mol^{-1} for gallic acid and β-glucogallin molecular weight, respectively.

*: Data expressed by gram of fresh weight (FW).

**: Data expressed by gram of dry weight (DW).

References: 1: Pyo *et al*., 2004, 2: Krenek *et al*., 2014, 3: Abbasi *et al*., 2015, 4: Palafox-Carlos *et al*., 2012, 5 : Romani *et al*., 2002, 6 : Tomaino *et al*.,2010, 7 : Rossetto *et al*., 2005, 8: Schmitzer *et al*., 2011, 9: Li *et al*., 1993, 10: Ayaz *et al*., 2008, 11 : Fang *et al*., 2008, 12: Zarena & Sankar, 2012, 13: Chen *et al*., 2008a, 14: Chen *et al*., 2008b, 15: Manpong *et al*., 2009, 17: Li *et al*.,1993,18: Tako *et al*., 2014, 19: Sanz *et al*., 2011, 20: De Vasconcelos *et al*., 2007, 21 : Fernández de Simón *et al*., 2006, 22 : Cadahía *et al*., 2001, 23 : Esmaeili *et al*., 2011,

24 : Kivilompolo *et al*., 2007, 25: Rani & Pratyusha, 2013, 26: Osman *et al*., 2004, 27: Veberic *et al*., 2008, 28: Del Verde-Méndez *et al*., 2003, 29 :Fang *et al*., 2007, 30: Santos *et al*., 2011, 31 : Boulekbache-Makhlouf *et al*., 2010, 32 : Chapuis-Lardy *et al*., 2002,34 : Shan *et al*., 2005,35:Pathak *et al*., 2004, 36: Bianco *et al*., 2003, 37: Khallouki *et al*., 2012, 38: Mertz *et al*., 2007, 39: Wada & Ou, 2002, 40: Määtä-Riihinen *et al*., 2004, 41 : Hoang *et al*., 2007, 42: Chen *et al*., 2008, 43: Sanz *et al*., 2010, 44: Loizzo *et al*., 2015, 45 : Gorinstein *et al*., 2004, 46 : Hu *et al*., 2009, 47 : Luo *et al*., 2002, 48 : Kowalski & Kowalska, 2005, 49: Reddivari *et al*., 2007, 50: Onyeneho & Hettiarachchy, 1993, 51: Chen *et al*., 2008, 52: Singh *et al*., 2014, 53: Jiang *et al*., 2013.

Supplemental Table 2.

Species (Family)	Molecule	Organ	Reference
Mangifera indica (Anacardiaceae)	Galloyl diglucoside GA-glucose ether	Fruit/pulp	Krenek *et al*., 2014
Robinia pseudoacacia (Fabaceae)	Gallic aldehyde Methyl-gallate	Heartwood	Sanz *et al*., 2011
Eucalyptus globulus (Myrtaceae)	Methyl gallate Digalloyl glucose	Bark	Santos *et al*., 2011
	methyl-digalloylglucose digalloylglucose	Flower	Boulekbache-Makhlouf *et al*., 2010
Camellia sinensis (Theaceae)	Galloylquinic acid	Buds, leaves stems	Jiang *et al*., 2013
Pistacia lentiscus (Anacardiaceae)	5-*O*-galloylquinic acid 3,5-*O*-digalloylquinic acid 3,4,5-*O*-trigalloylquinic acid	Leaves	Romani *et al*., 2002
Myrothamnus flabellifolius (Myrothamnaceae)	3,4,5-*O*-trigalloylquinic acid	Leaves	Moore *et al*., 2005
Manilkara zapota (Sapotaceae)	methyl 4-*O*-galloylchlorogenate 4-*O*-galloylchlorogenic acid	Fruit	Ma *et al*., 2003
Cornus officinalis (Cornaceae)	7-*O*-galloyl-D-sedoheptulose 7-*O*-galloylsecologanol (cornuside)	Fruit	Park *et al*., 2012
Mallotus japonicus (Euphorbiaceae)	Galloylbergenin Galloylnorbegenin Galloylglycerol Galloylshikimic acid	Bark	Saijo *et al*., 1990
Schinus terebinthifolius (Anacardiaceae)	Galloylshikimic acids	Fruit exocarp	Feuereisen *et al*., 2014
Inga laurina (Fabaceae)	Galloyl tyrosine	Young leaves	Lokvam *et al*., 2007
Acer sp. (Aceraceae)	Galloyl cyanidin glucosides	Red leaves	Ji *et al*., 1992
Victoria amazonica (Nymphaeceae)	Anthocyanidin 2"-*O*-galloylgalactosides		Strack *et al*., 1992 Fossen *et al*., 1998 Fossen & Andersen, 1999a
Acer platanoides (Aceraceae)	Cyanidin 3-(2", 3"-digalloylglucoside)	Red leaves	Fossen & Andersen, 1999b
Acalypha hispida (Euphorbiaceae)	Galloylated anthocyanins	Whole flowers	Reiersen *et al*., 2003
Acer okamotoamum (Aceraceae)	Flavonol digalloylglucoside	Leaves	Kim *et al*., 1998
Loropetalum chinense (Hammamelidaceae)	Flavonol digalloylglucoside	Leaves	Liu *et al*., 1997
Pemphis acidula (Lythraceae)	Quercetin / kaempferol 6"-*O*-galloylglycosides	Leaves	Masuda *et al*., 2001
Nymphaea candida (Nymphaeceae)	kaempferol 3-*O*-(2"-*O*-galloylrutinoside)	Flower	Liu *et al*., 2007
Rumex acetosa (Polygonaceae)	EC**G** Epiafzelechin-(4β→8)-EC**G** Epiafzelechin-(4β→6)-EC**G** Epiafzelechin-3-*O*-**G**-(4β→8)-EC**G** B2-3'-*O*-**G** B2-3,3'-di-*O*-**G** B5-3'-*O*-**G** B5-3,3'-di-*O*-**G** Cinnamtannin B1-3-*O*-**G**	Aerial parts	Bicker *et al*., 2009
Theobroma cacao (Sterculiaceae)	EC**G**	Shoot, young leaves	Osman *et al*., 2004
Fagopyrum esculentum (Polygonaceae)	EC-(4β→8)-EC-3-*O*-(3,4)-dimethyl-**G** Epiafzelechin-(4β→8)-EC-3-*O*-(3,4)-dimethyl-**G** Epiafzelechin-(4β→8)-EC-3-*O*-4-methyl-**G**	Grain	Ölschläger *et al*., 2008

	Epiafzelechin-(4β→8)-epiafzelechin-(4β→8)-EC-3-*O*-(3,4)-dimethyl-**G**		
Cornus officinalis (Cornaceae)	E(G)C 3-*O*-**G**-4-hydroxyethyl thioether	Fruit	Park *et al*., 2012
Pyrus communis (Rosaceae)	Monogalloylated procyanidin oligomer	Fruit	Es-Safi *et al*., 2006
Myrica rubra (Myricaceae)	EGC**G**	Leaves	Yang *et al*., 2011
Myrothamnus flabellifolia (Hamamelidaceae)	B4-3'-*O*-**G** B2-3'-*O*-**G** B2-3,3'-di-*O*-**G** B5-3,3'-*O*-**G** catechin-(4α→8)-EGC**G** C1-3"-*O*-**G** EC**G**-(4β→6)-EC-3-*O*-*p*-hydroxybenzoate	Twig tips	Anke *et al*., 2008
Burkea africana (Caesalpiniaceae)	(-)-fisetinidol-(4α→8)-catechin-3-*O*-**G**	Heartwood	Malan *et al*., 1988
Stryphnodendron adstringens (Fabaceae)	Robinetinidol-(4β→8)-EGC**G** Robinetinidol-(4α→8)-EGC**G**	Stem bark	de Mello *et al*., 1996
Gossypium hirsutum (Malvaceae)	Galloylated PAs	Cotton fibre	Feng *et al*., 2014
Hammamelis virginiata (Hamamelidaceae)	Galloylated PAs	Bark	Touriño *et al*., 2008
Anacardium occidentale (Anacardiaceae)	EGC**G**	Fruit	Gordon *et al*., 2012
Tea sinensis (Theaceae)	EC**G** EGC**G**	Bud, young leaves, stem	Jiang *et al*., 2013
Diospyros kaki (Ebenaceae)	(E)GC**G** (E)C**G**	Pulp	Li *et al*., 2010

Supplemental Table 2. Examples of galloylated derivatives identified in plants.
Abbreviations: Monomers: **G**: gallate, EC: Epicatechin, ECG : Epicatechin-3-*O*-gallate, E(G)C: Epi(gallo)catechin, EGCG: Epigallocatechin 3-*O*-gallate.
Dimers: B2: epicatechin-(4β→8)-epicatechin, B4: catechin-(4α→8)-epicatechin, B5: epicatechin-(4β→6)-epicatechin.
Trimers: C1: epicatechin-(4β→8)-epicatechin-(4β→8)-epicatechin.

Supplemental Table 3.

Purpose	Primer name	5'-3' sequence	Forward (F) or reverse (R)
qPCR tobacco leaves	PP2AtabacQF	GACCCTGATGTTGATGTTCGCT	F
	PP2AtabacQR	GAGGGATTTGAAGAGAGATTTC	R
	GAT1tabacQF ter	TATTGGTCGAGGTGCAACAA	F
	GAT1tabacQR ter	CGGGTATGGAACAACCAAGT	R
	GAT2tabacQF bis	GCGACCATGGTATGTTGATG	F
	GAT2tabacQR bis	CAACCAAGGCAAGAGATTCC	R
Cloning in pCR/GW/TOPO	GAT1TopoPC8F	GCCACCATGGCTGCAAAAAGGCCATCCAATTC	F
	GAT1TopoR	CTAGAGAGGGTAGTAATGCACCCACC	R
	GAT2TopoPC8F	GCCACCATGGCAACGGCCACACCTTCTCAAGAAGAGC	F
	GAT2TopoR	TTAGATGGGATAATAAGCAAACCACCTA	R
pEAQ vectors	pEAQF	CGTGGTTTTCGAACTTGGAG	F
	pEAQR	GAAAACCGCTCACCAAACAT	R

Supplemental Table 3. List of primers used for VvGAT study.

Supplemental Table 4.

Species	Common name	Family	ID (Accession n°)	Group
Arabidopsis thaliana	Arabidopsis	Brassicaceae	AtSDH (AAF08579)	I
Arabidopsis lyrata	-	Brassicaceae	AlSDH (EFH60832.1)	I
Brachypodium distachyon	-	**Poaceae**	**BdSDH1 (Bradi2g13170.3)**	**M-IV**
			BdSDH2 (Bradi4g05897.1)	**M-II**
Camellia sinensis	Tea plant	Theaceae	CasSDH1 (AIZ93902)	II
			CasSDH2 (AJA40947)	III
			CasSDH3 (AJA40948)	IV
Carica papaya	Papaya	Caricaceae	CpSDH (evm.model.supercontig_16.98)	I
Citrus clementina	Clementine	Rutaceae	CcSDH1 (Ciclev10004711m)	I
			CcSDH2 (Ciclev10000874m)	II
			CcSDH3 (Ciclev10020342m)	V
Citrus sinensis	Orange	Rutaceae	CsSDH1 (orange1.1g010050m)	II
			CsSDH2 (orange1.1g010101m)	V
			CsSDH3 (orange1.1g007151m)	I
Diospyros kaki	Persimmon	Ebenaceae	DkSDH1 (BAI40147)	IV
Elaeis guinensis	**African oil palm**	**Arecaceae**	**ElgSDH1 (XP_010914407.1)**	**M-I**
			ElgSDH2 (XP_010914405.1)	**M-I**
Eucalyptus grandis	Eucalyptus	Myrtaceae	EgSDH1 (Eucgr.H01214.1)	V
			EgSDH2 (Eucgr.H04428.1)	IV
			EgSDH3 (Eucgr.H04427.1)	III
			EgSDH4 (Eucgr.B01770.2)	II
			EgSDH5 (Eucgr.J00263.6)	I
Fagus sylvatica	Beech	Fagaceae	FsSDH (ABA54867.1)	I
Fragaria vesca	Strawberry	Rosaceae	FvSDH1 (XP_004302480)	III
			FvSDH2 (XP_004302479)	IV
			FvSDH3 (XP_004289250)	-
			FvSDH4 (XP_004288087)	I
Glycine max	Soybean	Fabaceae	GmSDH1 (XP_006606496)	IV
			GmSDH2 (XP_003520825)	II
			GmSDH3 (XP_006573239)	I
Gossypium raimondii	Cotton plant	Malvaceae	GrSDH1 (Gorai.006G075100)	I
			GrSDH2 (Gorai.013G116200)	II
			GrSDH3 (Gorai.004G250400)	IV
			GrSDH4 (Gorai.007G132000)	V
Jatropha curcas	Jatropha	Euphorbiaceae	JcSDH1 (XP_012089153.1)	II
			JcSDH2 (KDP40034.1)	I
			JcSDH3 (KDP23584.1)	III
			JcSDH4 (KDP23583.1)	IV
			JcSDH5 (KDP21552.1)	V
Juglans regia	Walnut	Juglandaceae	JrSDH (AAW65140)	I
Linum usitatissumum	Flax	Linaceae	LuSDH1 (Lus10033981)	I
			LuSDH2 (Lus10012797)	I
Malus domestica	Apple	Rosaceae	MdSDH1 (MDP0000176374)	IV
			MdSDH2 (MDP0000170005)	I
			MdSDH3 (MDP0000279446)	I
			MdSDH4 (MDP0000193646)	IV
			MdSDH5 (MDP0000306786)	I
Manihot esculenta	Cassava	Euphorbiaceae	MeSDH1 (cassava4.1_005332m)	IV
			MeSDH2 (cassava4.1_030256m)	I
			MeSDH3 (cassava4.1_004545m)	III
			MeSDH4 (cassava4.1_008122m)	V
Medicago truncatula	Barrel medic	Fabaceae	MtSDH (XP_003608198)	I
Musa accuminata	**Banana**	**Musaceae**	**MaSDH1 (XP_009386288.1)**	**M-I**
			MaSDH2 (XP_009408239.1)	**M-I**
			MaSDH3 (XP_009408237.1)	-
			MaSDH4 (XP_009408236.1)	-
Nelumbo nucifera	Lotus	Nelumbonaceae	NnSDH1 (XP_010244660.1)	I
			NnSDH2 (XP_010240930.1)	I
Nicotiana tabacum	Tobacco	Solanaceae	NtSDH1 (AAS90325)	I
			NtSDH2 (AAS90324)	II
Oryza sativa	**Rice**	**Poaceae**	**OsSDH1 (LOC_Os12g34874.2)**	**M-II**
			OsSDH2 (LOC_Os01g27750.1)	**M-IV**
Panicum hallii	**Hall's panicgrass**	**Poaceae**	**PhSDH1 (Pahal.0129s0053.1)**	**M-III**
			PhSDH2 (Pahal.0129s0052.1)	**M-II**
			PhSDH3 (Pahal.0012s0288.1)	**M-IV**
Panicum virgatum	**Switchgrass**	**Poaceae**	**PavSDH1 (Pavir.Cb00904.1)**	**M-III**
			PavSDH2 (Pavir.J24113.1)	**M-II**
			PavSDH3 (Pavir.J15887.1)	**M-III**
			PavSDH4 (Pavir.J01954.1)	**M-II**

Phaseolus vulgaris	Common bean	Fabaceae	PvSDH1 (Phvul.006G096400)	II
			PvSDH2 (Phvul.007G183800)	IV
			PvSDH3 (Phvul.003G035000)	I
Phoenix dactylifera	**Date palm**	**Arecaceae**	**PdSDH1 (XP_008813743.1)**	**M-I**
			PdSDH2 (XP_008813741.1)	**M-I**
Populus euphratica	Desert poplar	Salicaceae	PeSDH1 (XP_011010977.1)	I
			PeSDH2 (XP_011048878.1)	II
			PeSDH3 (XP_011024056.1)	IV
Populus trichocarpa	Poplar	Salicaceae	Poptr1 (Potri.010G019000)	I
			Poptr2 (Potri.013G029900)	II
			Poptr3 (Potri.005G043400)	II
			Poptr4 (Potri.014G135500)	V
			Poptr5 (Potri.013G029800)	IV
Prunus mume	Chinese plum	Rosaceae	PmSDH1 (XP_008244632.1)	I
			PmSDH2 (XP_008244628.1)	I
			PmSDH3 (XP_008246555.1)	IV
Prunus persica	Peach	Rosaceae	PpSDH1 (EMJ08566)	IV
			PpSDH2 (EMJ08564)	I
			PpSDH3 (EMJ06228)	I
Pyrus x bretschneideri	Chinese white pear	Rosaceae	PbSDH1 (XP_009341996)	IV
			PbSDH2 (XP_009333763)	I
			PbSDH3 (XP_009333759.1)	I
			PbSDH4 (XP_009373929.1)	IV
Ricinus communis	Castor oil plant	Euphorbiaceae	RcSDH1 (EEF49751)	III
			RcSDH2 (EEF49750)	II
			RcSDH3 (EEF45470)	I
			RcSDH4 (EEF41656)	V
Salix purpurea	Purple willow	Salicaceae	SpSDH1 (SapurV1A.0511s0170.1)	II
			SpSDH2 (SapurV1A.0353s0030.1)	IV
			SpSDH3 (SapurV1A.0053s0210.1)	V
			SpSDH4 (SapurV1A.0656s0090.1)	I
			SpSDH5 (SapurV1A.0353s0010.1)	II
			SpSDH6 (SapurV1A.0353s0020.1)	III
Sesamum indicum	Sesame	Pedaliaceae	SiSDH1 (XP_011079751.1)	V
			SiSDH2 (XP_011078708.1)	II
			SiSDH3 (XP_011071496.1)	I
Setaria italica	**Foxtail millet**	**Poaceae**	**SeiSDH1 (Si021763m)**	**M-III**
			SeiSDH2 (Si021701m)	**M-II**
			SeiSDH3 (Si000995m)	**M-IV**
Solanum lycopersicum	Tomato	Solanaceae	SlSDH1 (AAC17991)	I
			SlSDH2 (XP_010327280)	II
			SlSDH3 (XP_004242317)	V
Solanum tuberosum	Potato	Solanaceae	StSDH1 (XP_006352799.1)	V
			StSDH2 (XP_006342730.1)	I
			StSDH3 (XP_006338901.1)	II
Sorghum bicolor	**Sorghum**	**Poaceae**	**SbSDH1 (Sobic.008G116100.1)**	**M-III**
			SbSDH2 (Sobic.008G116000.1)	**M-II**
			SbSDH3 (Sobic.003G169900.1)	**M-IV**
Theobroma cacao	Cacao tree	Malvaceae	TcSDH1 (XP_007030118)	II
			TcSDH2 (EOY06376)	I
			TcSDH3 (EOY10614)	IV
			TcSDH4 (XP_007030116)	III
Triticum aestivum	**Common wheat**	**Poaceae**	**TaSDH1 (Traes_5AS_908BF5D38.1)**	**M-II**
			TaSDH2 (Traes_5DS_F55FF8108.1)	**M-II**
			TaSDH3 (Traes_3DL_2BA54361D.2)	**M-IV**
Vitis vinifera	Grapevine	Vitaceae	VvSDH5	I
			VvSDH7	II
			VvSDH8	III
			VvSDH9	IV
Zea mays	**Maize**	**Poaceae**	**ZmSDH1 (GRMZM5G804881_T01)**	**M-III**
			ZmSDH2 (GRMZM2G014376_T01)	**M-II**
			ZmSDH3 (GRMZM2G314652_T02)	**M-IV**

Supplemental Table 4. Plant DQD/SDHs used in this study.
Monocots sequences are shown in bold letters.

Supplemental Table 5.

Name	Len	cTP	mTP	SP	other	Loc	RC
PdSDH1	423	0.112	0.068	0.073	0.852	_	2
PdSDH2	516	0.016	0.061	0.855	0.294	S	3
ElgSDH1	513	0.016	0.059	0.862	0.285	S	3
ElgSDH2	513	0.122	0.133	0.154	0.687	_	3
TaSDH1	443	0.092	0.062	0.100	0.903	_	1
TaSDH2	475	0.051	0.291	0.069	0.847	_	3
TaSDH3	431	0.088	0.046	0.138	0.904	_	2
PhSDH1	527	0.436	0.125	0.419	0.066	C	5
PhSDH2	596	0.479	0.175	0.001	0.282	C	5
PhSDH3	509	0.031	0.024	0.881	0.185	S	2
PavSDH1	527	0.418	0.288	0.130	0.079	C	5
PavSDH2	498	0.263	0.099	0.276	0.268	S	5
PavSDH3	458	0.516	0.316	0.102	0.098	C	4
PavSDH4	542	0.717	0.123	0.253	0.020	C	3
SbSDH1	527	0.337	0.576	0.061	0.076	M	4
SbSDH2	541	0.703	0.120	0.249	0.037	C	3
SbSDH3	509	0.064	0.035	0.665	0.369	S	4
OsSDH1	547	0.731	0.177	0.099	0.040	C	3
OsSDH2	509	0.025	0.054	0.859	0.159	S	2
BdSDH1	509	0.082	0.032	0.701	0.447	S	4
BdSDH2	545	0.454	0.323	0.129	0.037	C	5
SeiSDH1	520	0.080	0.736	0.075	0.068	M	2
SeiSDH2	539	0.814	0.112	0.230	0.015	C	3
SeiSDH3	527	0.031	0.020	0.906	0.233	S	2
ZmSDH1	532	0.307	0.152	0.101	0.402	_	5
ZmSDH2	543	0.754	0.118	0.150	0.090	C	2
ZmSDH3	509	0.084	0.038	0.547	0.507	S	5
MaSDH1	511	0.021	0.082	0.887	0.178	S	2
MaSDH2	510	0.033	0.120	0.780	0.181	S	3
MaSDH3	452	0.099	0.040	0.061	0.863	_	2
MaSDH4	555	0.074	0.779	0.020	0.022	M	2
VvSDH5	519	0.145	0.067	0.142	0.895	_	2
VvSDH7	519	0.086	0.091	0.136	0.922	_	2
VvSDH8	531	0.569	0.063	0.100	0.419	C	5
VvSDH9	531	0.492	0.061	0.170	0.547	_	5
Poptr1	541	0.485	0.036	0.412	0.185	C	5
Poptr2	518	0.057	0.297	0.282	0.292	M	5
Poptr3	518	0.032	0.108	0.600	0.336	S	4
Poptr4	527	0.179	0.055	0.201	0.437	_	4
Poptr5	521	0.123	0.200	0.089	0.818	_	2
EgSDH1	585	0.693	0.229	0.003	0.169	C	3
EgSDH2	575	0.358	0.104	0.109	0.394	_	5
EgSDH3	530	0.644	0.062	0.077	0.476	C	5
EgSDH4	457	0.098	0.051	0.387	0.603	_	4
EgSDH5	521	0.061	0.084	0.376	0.701	_	4
TcSDH1	520	0.075	0.142	0.168	0.725	_	3
TcSDH2	563	0.116	0.077	0.155	0.774	_	2
TcSDH3	533	0.217	0.057	0.126	0.835	_	2
TcSDH4	475	0.143	0.465	0.034	0.295	M	5
MdSDH1	532	0.233	0.105	0.130	0.685	_	3
MdSDH2	602	0.139	0.051	0.155	0.773	_	2
MdSDH3	618	0.640	0.255	0.009	0.142	C	4
MdSDH4	425	0.139	0.059	0.102	0.907	_	2
MdSDH5	736	0.008	0.833	0.002	0.105	M	2
PvSDH1	505	0.074	0.159	0.093	0.620	_	3
PvSDH2	524	0.160	0.267	0.082	0.576	_	4
PvSDH3	526	0.120	0.059	0.211	0.840	_	2
CasSDH1	519	0.130	0.094	0.223	0.669	_	3
CasSDH2	532	0.800	0.037	0.164	0.302	C	3
CasSDH3	508	0.191	0.111	0.366	0.527	_	5
NtSDH2	521	0.032	0.060	0.773	0.243	S	3
NtSDH1	569	0.063	0.083	0.126	0.911	_	2
AtSDH	603	0.950	0.375	0.002	0.006	C	3
StSDH1	514	0.051	0.107	0.118	0.809	_	2
StSDH2	515	0.080	0.084	0.244	0.802	_	3
StSDH3	518	0.043	0.041	0.641	0.569	S	5
CsSDH1	519	0.038	0.084	0.591	0.538	S	5
CsSDH2	518	0.049	0.093	0.282	0.700	_	3
CsSDH3	616	0.253	0.078	0.127	0.600	_	4
GmSDH1	533	0.701	0.099	0.081	0.239	C	3
GmSDH2	524	0.008	0.141	0.617	0.355	S	4
GmSDH3	530	0.512	0.052	0.154	0.681	_	5

PpSDH1	502	0.116	0.080	0.129	0.837	_	2
PpSDH2	505	0.143	0.055	0.116	0.910	_	2
PpSDH3	535	0.240	0.194	0.110	0.302	_	5
RcSDH1	524	0.284	0.075	0.182	0.486	_	4
RcSDH2	519	0.078	0.056	0.369	0.706	_	4
RcSDH3	539	0.484	0.049	0.123	0.718	_	4
RcSDH4	515	0.161	0.096	0.273	0.468	_	5
FvSDH1	530	0.314	0.055	0.096	0.747	_	3
FvSDH2	534	0.395	0.098	0.157	0.373	C	5
FvSDH3	514	0.140	0.107	0.338	0.620	_	4
FvSDH4	531	0.313	0.174	0.059	0.583	_	4
DkSDH	508	0.321	0.119	0.116	0.701	_	4
PbSDH1	532	0.157	0.151	0.102	0.734	_	3
PbSDH3	537	0.062	0.170	0.162	0.757	_	3
PbSDH4	532	0.268	0.073	0.138	0.759	_	3
PbSDH2	541	0.477	0.068	0.205	0.616	_	5
GrSDH1	536	0.111	0.071	0.178	0.789	_	2
GrSDH2	509	0.124	0.109	0.178	0.788	_	2
GrSDH3	532	0.261	0.061	0.138	0.837	_	3
GrSDH4	519	0.059	0.044	0.290	0.844	_	3
MeSDH1	531	0.201	0.046	0.108	0.869	_	2
MeSDH2	517	0.174	0.067	0.162	0.804	_	2
MeSDH3	569	0.259	0.771	0.015	0.010	M	3
MeSDH4	429	0.050	0.081	0.228	0.784	_	3
AlSDH	603	0.967	0.157	0.002	0.017	C	1
MtSDH	506	0.090	0.042	0.518	0.650	_	5
SlSDH1	545	0.039	0.166	0.105	0.926	_	2
SlSDH2	514	0.034	0.037	0.819	0.424	S	4
SlSDH3	532	0.021	0.431	0.032	0.871	_	3
FsSDH	536	0.297	0.057	0.303	0.458	_	5
JrSDH	508	0.171	0.080	0.317	0.602	_	4
PmSDH2	533	0.159	0.111	0.170	0.742	_	3
PmSDH1	536	0.159	0.111	0.170	0.742	_	3
PmSDH3	526	0.298	0.277	0.052	0.489	_	5
PeSDH1	538	0.400	0.045	0.444	0.180	S	5
PeSDH2	518	0.039	0.091	0.696	0.261	S	3
PeSDH3	521	0.101	0.236	0.109	0.715	_	3
NnSDH1	358	0.022	0.047	0.732	0.570	S	5
NnSDH2	527	0.079	0.090	0.451	0.670	_	4
CcSDH1	531	0.253	0.078	0.127	0.600	_	4
CcSDH2	519	0.038	0.084	0.591	0.538	S	5
CcSDH3	417	0.056	0.090	0.086	0.920	_	1
CpSDH	464	0.097	0.097	0.179	0.838	_	2
LuSDH1	534	0.186	0.037	0.176	0.761	_	3
LuSDH2	534	0.218	0.030	0.170	0.752	_	3
SpSDH1	519	0.120	0.176	0.444	0.196	S	4
SpSDH2	533	0.146	0.127	0.093	0.712	_	3
SpSDH3	515	0.203	0.051	0.282	0.438	_	5
SpSDH4	541	0.315	0.051	0.346	0.417	_	5
SpSDH5	532	0.040	0.542	0.132	0.342	M	5
SpSDH6	446	0.627	0.096	0.057	0.436	C	5
SiSDH1	470	0.044	0.376	0.033	0.803	_	3
SiSDH2	517	0.060	0.256	0.178	0.453	_	5
SiSDH3	517	0.031	0.165	0.283	0.659	_	4
JcSDH1	519	0.070	0.120	0.500	0.249	S	4
JcSDH2	539	0.213	0.083	0.230	0.581	_	4
JcSDH3	531	0.209	0.143	0.089	0.580	_	4
JcSDH4	531	0.157	0.179	0.053	0.737	_	3
JcSDH5	519	0.072	0.097	0.113	0.787	_	2

Supplemental Table 5.TargetP v1.1 prediction results for plant shikimate dehydrogenases (using PLANT networks).
Len: Sequence length; **cTP, mTP, SP, other**: Final NN scores on which the final prediction is based (Loc, see below). Note that the scores are not really probabilities, and they do not necessarily add to one. However, the location with the highest score is the most likely according to TargetP, and the relationship between the scores (the reliability class, see below) may be an indication of how certain the prediction is; **C**: Chloroplast, i.e. the sequence contains cTP, a chloroplast transit peptide; **M**: Mitochondrion, i.e. the sequence contains mTP, a mitochondrial targeting peptide; **S**: Secretory pathway, i.e. the sequence contains SP, a signal peptide; **_**: Any other location; **RC**: Reliability class, from 1 to 5, where 1 indicates the strongest prediction. RC is a measure of the size of the difference ('diff') between the highest (winning) and the second highest output scores. There are 5 reliability classes, defined as follows: 1: diff > 0.800, 2: 0.800 > diff > 0.600, 3: 0.600 > diff > 0.400, 4: 0.400 > diff > 0.200, 5: 0.200 > diff. Thus, the lower the RC value, the safer the prediction.

Supplemental Table 6.

Name	Len	cTP	mTP	SP	other	Loc	RC
GAT1	470	0.008	0.307	0.744	0.028	S	3
GAT2	464	0.096	0.079	0.687	0.173	S	3
VvSCP3	444	0.009	0.038	0.858	0.156	S	2
VvSCP4	437	0.036	0.067	0.527	0.439	S	5
VvSCP5	464	0.038	0.074	0.862	0.123	S	2
VvSCP7	439	0.026	0.025	0.809	0.037	S	2
VvSCP8	501	0.027	0.050	0.943	0.040	S	1
VvSCP10	468	0.003	0.367	0.697	0.114	S	4
VvSCP13	453	0.219	0.069	0.476	0.323	S	5
VvSCP14	470	0.013	0.131	0.874	0.029	S	2
VvSCP15	469	0.009	0.129	0.891	0.024	S	2
VvSCP16	482	0.013	0.058	0.846	0.102	S	2
VvSCP18	501	0.350	0.021	0.704	0.003	S	4
VvSCP19	452	0.011	0.020	0.928	0.030	S	1
VvSCP20	504	0.005	0.026	0.991	0.037	S	1
VvSCP21	472	0.005	0.023	0.978	0.053	S	1
VvSCP24	462	0.006	0.118	0.507	0.438	S	5
VvSCP25	495	0.001	0.169	0.934	0.089	S	2
VvSCP26	550	0.010	0.135	0.101	0.875	_	2
VvSCP27	461	0.012	0.193	0.915	0.018	S	2
VvSCP29	488	0.012	0.156	0.895	0.015	S	2
VvSCP31	469	0.007	0.072	0.979	0.008	S	1
VvSCP32	534	0.008	0.078	0.733	0.409	S	4
VvSCP33	488	0.011	0.115	0.947	0.020	S	1
VvSCP34	501	0.191	0.033	0.922	0.001	S	2
VvSCP35	469	0.010	0.423	0.312	0.085	M	5
VvSCP36	463	0.006	0.035	0.961	0.056	S	1
VvSCP37	451	0.016	0.054	0.965	0.021	S	1
VvSCP38	472	0.003	0.026	0.944	0.202	S	2
VvSCP39	456	0.011	0.008	0.966	0.095	S	1
VvSCP40	497	0.012	0.093	0.982	0.002	S	1
VvSCP42	474	0.003	0.093	0.881	0.252	S	2
VvSCP44	476	0.005	0.022	0.994	0.028	S	1
VvSCP45	461	0.012	0.200	0.906	0.019	S	2
VvSCP46	448	0.111	0.290	0.006	0.872	_	3
VvSCP47	504	0.175	0.008	0.926	0.041	S	2
VvSCP48	382	0.018	0.035	0.969	0.018	S	1
VvSCP50	473	0.004	0.011	0.986	0.179	S	1
VvSCP51	481	0.003	0.311	0.950	0.020	S	2
VvSCP52	461	0.003	0.121	0.936	0.093	S	1
VvSCP53	456	0.005	0.069	0.964	0.043	S	1
VvSCP54	500	0.005	0.892	0.168	0.105	M	2
VvSCP55	463	0.005	0.068	0.922	0.043	S	1
VvSCP56	480	0.014	0.061	0.947	0.033	S	1
VvSCP57	472	0.001	0.466	0.852	0.027	S	4
VvSCP58	479	0.004	0.012	0.987	0.078	S	1
VvSCP59	449	0.047	0.044	0.865	0.101	S	2
VvSCP60	491	0.097	0.514	0.082	0.316	M	5
VvSCP61	467	0.005	0.074	0.989	0.017	S	1
VvSCP62	482	0.087	0.098	0.266	0.607	_	4
VvSCP63	458	0.014	0.033	0.873	0.064	S	1
SlGAC	464	0.011	0.085	0.960	0.051	S	1
CtAT1	469	0.004	0.188	0.976	0.008	S	2
AtSMT	433	0.010	0.036	0.978	0.059	S	1
AtSCT	464	0.004	0.051	0.991	0.014	S	1
BnSCT1	466	0.007	0.057	0.981	0.026	S	1
BnSCT2	466	0.007	0.055	0.981	0.028	S	1
AtSAT	437	0.005	0.049	0.983	0.036	S	1
AtSST	437	0.018	0.106	0.877	0.073	S	2
AsSCPL1	493	0.012	0.100	0.953	0.030	S	1
DkSCPL1	452	0.023	0.056	0.451	0.821	_	4
DkSCPL2	491	0.007	0.570	0.816	0.010	S	4
AtSCPL17	437	0.002	0.061	0.982	0.061	S	1
CsSCPL	480	0.004	0.552	0.672	0.271	S	5
SbHNL1	510	0.492	0.445	0.037	0.090	C	5
TaCPDWII	444	0.035	0.145	0.103	0.880	_	2
HvSCPI	499	0.002	0.684	0.777	0.009	S	5
HvCPDWII	436	0.010	0.476	0.202	0.577	_	5
HvSCPIII	508	0.020	0.118	0.821	0.004	S	2
ScCPY	532	0.046	0.241	0.676	0.075	S	3
BRS1	465	0.294	0.074	0.234	0.040	C	5
SCPLe	498	0.002	0.131	0.988	0.018	S	1

MtSCP1	495	0.006	0.044	0.981	0.019	S	1
DgSCPL1	475	0.007	0.043	0.982	0.027	S	1
DgSCPL2	470	0.003	0.087	0.986	0.012	S	1
DgSCPL3	456	0.001	0.046	0.929	0.177	S	2

Supplemental Table 6. TargetP v1.1 prediction results for Serine Carboxypeptidases (using PLANT networks).
Len: Sequence length; **cTP, mTP, SP, other**: Final NN scores on which the final prediction is based (Loc, see below). Note that the scores are not really probabilities, and they do not necessarily add to one. However, the location with the highest score is the most likely according to TargetP, and the relationship between the scores (the reliability class, see below) may be an indication of how certain the prediction is; **C**: Chloroplast, i.e. the sequence contains cTP, a chloroplast transit peptide; **M**: Mitochondrion, i.e. the sequence contains mTP, a mitochondrial targeting peptide; **S**: Secretory pathway, i.e. the sequence contains SP, a signal peptide; **_**: Any other location; **RC**: Reliability class, from 1 to 5, where 1 indicates the strongest prediction. RC is a measure of the size of the difference ('diff') between the highest (winning) and the second highest output scores. There are 5 reliability classes, defined as follows: 1: diff > 0.800, 2: 0.800 > diff > 0.600, 3: 0.600 > diff > 0.400, 4: 0.400 > diff > 0.200, 5: 0.200 > diff. Thus, the lower the RC value, the safer the prediction.