

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE MONTPELLIER SUPAGRO

En Microbiologie / Bio-informatique

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau

Portée par

Unité de recherche Biologie et Génétique des Interactions Plante-Parasite

Influence de l'hôte et de phytovirus sur l'assemblage du microbiote du riz à l'échelle de deux agrosystèmes

Présentée par Pascal Alonso

Le 12 juin 2020

Sous la direction de Christian VERNIERE
et Philippe ROUMAGNAC

Devant le jury composé de

Wafa ACHOUAK, Dr. directrice de recherche, CNRS Aix-Marseille

Alain SARNIGUET, Dr. directeur de recherche, INRAE Angers

Claire NEEMA, Pr. L'institut Agro Montpellier

Sébastien MASSART, Pr., Université de Liège

Philippe ROUMAGNAC, Dr., CIRAD Montpellier

Christian VERNIERE, Dr., CIRAD Montpellier

Rapportrice

Rapporteur

Examinatrice

Examineur

Co-directeur de thèse

Directeur de thèse



UNIVERSITÉ
DE MONTPELLIER



Remerciements

Tout d'abord, je souhaite remercier Wafa ACHOUAK, Alain SARNIGUET, Claire NEEMA et Sébastien MASSART pour avoir accepté d'être membre de mon jury de thèse et d'évaluer mon travail. J'espère que vous prendrez autant de plaisir à lire ce manuscrit que j'en ai pris pour réaliser ce travail durant ces trois années. Ensuite je tiens à remercier le CIRAD, Agropolis Fondation et l'INRA qui ont été les financeurs de ce projet et m'ont permis de réaliser ce travail de thèse.

Ce manuscrit est le support de l'évaluation de mon travail de thèse, mais il est aussi le fruit des efforts conjoint de toute une équipe que je tiens à remercier pour sa participation, son aide et son soutien.

Ainsi, je souhaite remercier chaleureusement mes deux encadrants, Christian VERNIERE et Philippe ROUMAGNAC. Vous m'avez beaucoup apporté. Votre confiance, votre écoute, vos encouragements constants et votre disponibilité m'ont permis de m'approprier ce sujet et de m'épanouir tout au long de ce travail. Vous avez été pour moi d'excellents encadrants et j'espère avoir été à la hauteur de tout ce que vous m'avez apporté.

Christian, je te remercie pour ton accueil lors de notre première rencontre, il y a quatre ans. Cette rencontre aura été importante dans mes choix futurs. Tu m'as présenté un beau sujet et tu as réussi à me convaincre de me lancer dans l'exigent défi de la thèse. Merci pour ton support sans faille, et ce, dans toutes les épreuves de la thèse (à la fois par tes conseils, ton aide dans des manips sans fin et dans le difficile exercice de l'écriture). Ta sagesse dans l'encadrement m'a permis de rester sur les rails tout en ayant la possibilité de choisir la destination.

Philippe, merci pour ton encadrement, et pour ta capacité à fédérer les personnes. Ton assurance, tes encouragements, tes critiques toujours positives m'ont donné une grande confiance en moi. Je te remercie pour tes conseils et ton soutien à la fois, dans la réflexion de recherche, dans l'écriture et dans l'exercice des présentations orales.

Après avoir remercié pilote et copilote de l'équipe MicroQuar, je voudrais remercier toute l'équipe à bord. Vous m'avez tous aidé, encouragé, conseillé, écouté. C'est avec regret que j'écris ces remerciements, car je me rends compte qu'il est temps de descendre du bus MicroQuar et pourtant je m'y sentais drôlement bien. Entre nous, jetez un œil dans le coffre : je vous ai laissé de jolis camemberts rôtis à l'ail spécial ALONSO (les témoins du voyage pour Pathobiome vous expliqueront).

Bien évidemment, je remercie Laurence d'avoir été mon binôme tout au long de la thèse. Nous avons fait un beau travail d'équipe. Merci, pour tes encouragements qui ont sauvé quelques fois la vie de mon PC. Merci pour ton aide, grâce à toi je n'ai pas connu le sentiment de solitude du doctorant finissant ses manips à la fermeture du labo. Merci, pour les sacrifices de temps, en particulier pour les soirées microbio ou les weekends R, qui ont souvent remplacé tes sorties randonnées ou tes apéros entre amis. Merci, d'avoir pallié à mon syndrome d'oubli des prénoms et des dates. Je pourrais te remercier pour encore de nombreuses choses que tu m'as apportées

durant la thèse : je ferai cela par texto, ou lors d'une cérémonie officielle, une médaille (ou un pin's) te sera attribuée, car me supporter pendant trois ans...

Merci Charlotte et Lisa pour votre aide dans les prélèvements en rizières, pour les analyses du virome du riz, mais surtout pour les bons moments passés ensemble en particulier lors des dégustations Camargue 2018 du L50. Merci Serge et Cécile, en plus des bon moments votre aide m'a été précieuse lors des prélèvements en Camargue ou de l'opération décorticage de grains de riz. Merci Oumi et Pei-Jung, cela été un plaisir de participer à votre stage de master. Grâce à votre aide et votre efficacité, nous avons réussi à obtenir de précieux résultats. J'espère que cette expérience aura contribué à vous donner l'envie de poursuivre votre future carrière dans la recherche.

Manu, merci de m'avoir formé efficacement à l'analyse VANA. J'espère que j'aurai l'occasion d'apprendre plus de choses à toi sur les possibilités qu'offre le Minion dans les prochains mois. Merci Jean pour tes conseils en début de thèse.

La bio-informatique était un grand défi dans ma thèse. Merci à Denis, Fred et Hervé pour votre aide dans ce domaine. Vous avez tous les trois toujours été patient et d'une incroyable disponibilité. Merci Fred de m'avoir converti au Tidyverse, aujourd'hui je ne peux plus m'en passer.

Romain, merci ta bonne humeur, et pour avoir accepté d'échanger quelques kilos et un poumon contre des échantillons de riz en chine.

Merci Isabelle, pour avoir partagé les missions en Camargue, j'espère que nous pourrons valoriser ce travail prochainement et arriver peut-être à réunir archées et bactéries.

Florence et Agnès, merci pour votre soutien votre support sur les parties administratives où je n'excelle pas véritablement.

Plus généralement, je tenais à remercier les membres de l'unité BGPI, pour leurs accueils et leurs gentillesse dans les moments importants tant professionnels que personnels (premier séminaire, mariage, et naissance). Merci Pierre, pour toutes les tentatives d'analyse des résultats de génotypage des riz qui ont abouties à un joli résultat ; Merci François pour ton aide lorsque R me dépassait. Alexandre, Hilaire, j'ai beaucoup apprécié être votre collègue de bureau, j'ai beaucoup travaillé avec vous sur la thématique zygomatique. Merci Henri pour tes conseils jardins et ton partage de connaissances sur les maladies du riz. Je ne pourrais pas te donner la clé d'une infestation de Pyri réussie, mais, je pourrais t'expliquer comment organiser un BBQ au mas d'Adrien.

Je souhaite également remercier Eric et Frédérique de la plateforme Geneseq. J'espère pouvoir encore continuer à venir vous voir et collaborer avec vous.

Je voudrais aussi remercier Julien et Pierre du l'UMR AGAP, pour m'avoir formé à l'utilisation de la plateforme et pour tout le support que vous m'avez apporté par la suite. Mais aussi François Massol pour m'avoir pendant une semaine dédiée ton temps pour me former dans les

analyses de réseau d'interactions entre microorganismes. Je suis sûr que ces connaissances me serviront très prochainement.

Merci à Arnaud du Centre Français du Riz, qui a mis en place des conditions idéales pour réaliser nos travaux en Camargue et qui nous a permis d'en bénéficier.

J'adresse aussi mes remerciements à Jean-Christophe Avarre, pour m'avoir proposé à plusieurs reprises des opportunités qui ont été importantes dans ma carrière professionnelle.

Je remercie l'équipe de volley surtout les jeunes (Valentin, Jeremy, Thomas, Alex, Marlène) pour les pauses du mardi et du jeudi. Pour le tournoi, n'hésitez pas à rappeler encore et encore à Steph (et le reste des PCR) que nous sommes la SEULE équipe de Baillarguet à passer les poules.

Pour terminer, je remercie mes parents, qui m'ont toujours soutenu dans mon parcours. Je suis très fier de vous avoir. Je remercie aussi ma très très gentille et magnifique femme Sonia, pour ses encouragements à ne pas baisser les bras, son soutien quotidien, sa patience surtout dans ces dernières semaines, pour les relectures et corrections. Merci à ma petite merveille, Emma, d'avoir été patiente pendant cette période d'écriture. Merci Jordan et Guillaume pour les moments de pose pendant ces trois ans.

J'arrive au bout de cette expérience : Terminus Pascal descend... j'espère n'avoir oublié personne et j'espère pouvoir vous retrouver prochainement pour de nouvelles aventures.

Bonne continuation à vous tous

Résumé

Titres de la thèse : Influence de l'hôte et de phytovirus sur l'assemblage du microbiote du riz à l'échelle de deux agrosystèmes

Les communautés microbiennes (microbiote) des plantes sont de plus en plus prises en compte en santé des plantes car elles jouent un rôle actif dans la régulation des agents phytopathogènes. Toutefois, les processus d'assemblage du microbiote des plantes et des forces évolutives qui gouvernent ces processus restent loin d'être compris et maîtrisés.

Nous avons choisi d'aborder cette problématique en focalisant sur le modèle riz et en répondant aux objectifs suivants : (i) déterminer si la structure des communautés microbiennes associées au riz est influencée par le compartiment, le stade végétatif et le génotype de celui-ci et (ii) estimer si la structure des communautés bactériennes et fongiques est significativement modifiée par la présence/absence de phytovirus. Pour répondre à ces objectifs, nous avons caractérisé trois composantes majeures du microbiote des plantes, à savoir les virus, les bactéries et les champignons des racines et des tiges du riz dans deux agrosystèmes situés en Camargue (France) et dans les terrasses traditionnelles du Yuanyang (Chine).

La caractérisation du virome du riz a mis en évidence la présence d'une épidémie virale dans l'agrosystème chinois, pourtant présenté comme étant peu soumis aux bioagresseurs des plantes. Le phytovirus détecté dans 8,9% des plants échantillonnés est le Southern rice black-streaked dwarf virus (SRBSDV). La même analyse réalisée en France révèle une prévalence importante (11.6%) d'un endornavirus persistant du riz (*oryza sativa* alphaendornavirus, OsEV) qui a été retrouvé au sein de plantes asymptomatiques. Nous montrons par ailleurs que l'infection du riz à ces deux virus est génotype-dépendant.

La synthèse des résultats de caractérisation des communautés bactériennes et fongiques obtenus sur les deux sites d'étude met en lumière une dynamique dans le temps et dans l'espace de la structure des communautés microbiennes. Spécifiquement, nous montrons que (i) le type de tissu de la plante, aérien ou souterrain, interne ou externe, influence significativement la richesse et la composition des communautés microbiennes, (ii) le stade de développement du riz influence aussi significativement la composition des communautés microbiennes, (iii) le génotype du riz a en revanche un effet faible sur la structure des communautés microbiennes et enfin, (iv) la présence/absence des deux virus (SRBSDV et OsEV) ne modifie pas significativement la structure des communautés microbiennes du riz.

Nous montrons donc que l'assemblage du microbiote du riz est pour une partie le résultat de l'effet de l'hôte par des facteurs déterministes (sélection), mais que probablement d'autres facteurs écologiques de nature déterministe, par exemple à travers les interactions microbes-microbes, et de nature stochastique y contribuent participant à la diversité de l'holobionte (l'hôte et son microbiote) du riz.

Mots clés : Microbiote, holobionte, diversité, bactérie, champignon, virus, riz, génotype, métagénomique, communauté, interactions plante-microorganismes, dynamique

Abstract

Influence of host and phytovirus on rice microbiota assembly within two agrosystems

The microbial communities, or microbiota, associated with plants play an active role in plant health, i.e. they can regulate plant pathogens or promote plant growth. However, the assembly processes of the plant microbiota and the evolutionary forces that govern these processes remain far from being understood and controlled.

We have chosen to tackle this problem by focusing on the rice as a plant model and addressing the following objectives: (i) determining whether the structure of the rice-associated microbial communities is shaped by the plant compartment, the plant growth stage and the plant genotype, and (ii) estimating whether the structure of the bacterial and fungal communities is significantly influenced by the presence/absence of plant viruses. To fulfil these objectives, we characterized three major components of the plant-associated microbiota, i.e. viruses, bacteria and fungi from the roots and stems of rice in two agrosystems located in Camargue (France) and in the Honghe Hani traditional rice terraces system (China).

The characterization of the rice virome highlighted the presence of a plant viral outbreak in the Chinese agrosystem, though renowned for a limited impact of plant pathogens. The plant virus detected in 8.9% of the sampled plants was the Southern rice black-streaked dwarf virus (SRBSDV). A metagenomics-based approach was also carried out in France on asymptomatic plants and revealed a high prevalence (11.6%) of a persistent rice endornavirus (*oryza sativa* alphaendornavirus, OsEV). We show that both rice infections occurring in France and China were genotype-dependent.

In addition, the characterization of bacterial and fungal communities' at the French and Chinese study sites overall highlighted a temporal and spatial dynamic of the structure of the microbial communities. Specifically, we show that (i) the type of plant tissue, aboveground or belowground, internal or external, significantly influenced the richness and the composition of the microbial communities, (ii) the development stage of rice also influenced the composition of its microbial communities, (iii) the rice genotype had a weak effect on the microbial community structures and finally, (iv) the presence/absence of both plant viruses (SRBSDV and OsEV) did not significantly modify the rice microbial community structures.

We therefore showed that the assembly of the rice microbiota is partly the result of the effect of the host by deterministic factors (selection) but that probably other ecological factors with a deterministic nature, for example through microbe-microbe interactions, and stochastic nature, contribute to the rice microbiota assembly and further to the diversity of the rice holobiont (the host and its microbiota).

Keywords: Microbiota, holobiont, diversity, bacteria, fungi, virus, rice, genotype, metagenomics, community, plant-microorganism interactions, dynamic

Table des matières

Remerciements	1
Résumé	4
Abstract	6
Table des matières	7
Liste des figures et des tableaux.....	12
Liste des abréviations	16
Chapitre 1 : Synthèse Bibliographique.....	18
1.1. Etude des communautés microbiennes par des méthodes moléculaires	19
1.1.1. Méthodologie de la métagénomique ciblée.....	22
1.1.2. La métagénomique globale ou métagénomique shotgun	24
1.1.3. Séquençage de longues séquences nucléotidiques	26
1.2. Etude des communautés microbiennes par des méthodes de culture.....	27
1.3. Caractérisation de la diversité de l'hôte par typage génétique.....	29
1.4. Analyses bio-informatiques	30
1.4.1. Métagénomique ciblée	30
1.4.2. Métagénomique virale shotgun : analyses bio-informatiques des données de séquençages du virome par l'approche VANA.....	35
1.4.3. Analyses bio-informatiques des données de génotypage des plantes par GBS .	36
1.5. Limites des méthodes d'analyses des communautés microbiennes.....	37
1.5.1. Limites spécifiques à la métagénomique ciblée	37
1.5.2. Limites de la métagénomique globale.....	39
1.5.3. Limites des méthodes de culturomique	39
1.6. Estimation de la diversité des communautés microbiennes.....	40

1.6.1.	Indices de diversité	43
1.7.	De la notion de symbiose au concept de l'holobionte	47
1.8.	Le riz, une plante modèle	51
1.9.	Les communautés microbiennes associées au riz	52
1.9.1.	Les microbes colonisent différents compartiments de la plante	52
1.9.2.	Les communautés microbiennes des compartiments racinaires	54
1.9.3.	Les communautés microbiennes de la phyllosphère	58
1.9.4.	L'endosphère et les communautés microbiennes endophytes du riz	60
1.9.5.	Les communautés microbiennes de la graine	62
1.10.	Assemblage des communautés microbiennes du riz	63
1.10.1.	Origine du microbiote du riz	63
1.10.2.	Dynamique de l'assemblage du microbiote du riz	63
1.10.3.	Les facteurs influençant la composition et la structure du microbiote	65
1.11.	Etude du virome du riz	68
1.12.	Problématique et objectifs	69
Chapitre 2 : Caractérisation de l'holobionte du riz dans un agrosystème chinois centenaire et durable		72
2.1.	Introduction générale du chapitre 2	73
2.2.	Section 1 : Analyses du virome des variétés de riz traditionnelles et modernes cultivées au sein du village Malizhai	74
2.2.1.	Matériel et méthodes	74
2.2.2.	Résultats et discussion	77
2.2.3.	Conclusion	79
2.3.	Section 2: Emergence of Southern Rice Black-Streaked Dwarf Virus in the Centuries-Old Chinese Yuanyang Agrosystem of Rice Landraces	79

2.4.	Section 3: Heterogeneity of the rice microbial community of the Chinese centuries-old Honghe Hani rice terraces system	93
2.5.	Conclusions du chapitre 2	111
Chapitre 3: Etude de facteurs de structuration du microbiote endophyte du riz sur un site d'étude en Camargue.....		113
3.1.	Introduction.....	114
3.2.	Matériels et méthodes	116
3.2.1.	Système d'étude	116
3.2.2.	Echantillonnage	119
3.2.3.	Préparation des échantillonnages	120
3.2.4.	Analyse du virome par métagénomique VANA	122
3.2.5.	Analyse des communautés bactériennes par métagénomique ciblée	122
3.3.	Résultats et discussions.....	127
3.3.1.	Analyse du virome par métagénomique VANA	128
3.3.2.	Analyse du microbiote endophyte des racines, des tiges et des graines de riz en Camargue	132
3.3.3.	Conclusions	158
Discussion générale et perspectives		160
3.4.	Rappel des objectifs de thèse	161
3.5.	Caractérisation des microbiotes du riz	161
3.5.1.	Une dynamique dans l'espace	161
3.5.2.	Une dynamique dans le temps.....	162
3.5.3.	Influence du génotype de la plante hôte sur la structure des communautés microbiennes associées au riz	164
3.5.4.	Caractérisation du virome du riz et étude de son influence sur la composition des communautés microbiennes du riz.....	167

3.5.5. Processus déterministes et stochastiques façonnant les communautés microbiennes du riz.....	170
3.5.6. Vers une approche fonctionnelle du microbiote du riz ?	173
Références bibliographiques	176
Annexes	198

Liste des figures et des tableaux

Figures de la Synthèse bibliographique

Figure 1.1 : Structure secondaire de l'ARNr 16S d'*Escherichia coli* présentant les différentes régions variables de cette molécule

Figure 1.2 : Localisation génomique des espaceurs internes ITS1 et ITS2

Figure 1.3 : Illustration du recouvrement des résultats de la caractérisation des compositions des communautés microbiennes identifiées par une approche culture dépendante moderne et par une approche de pyroséquençage

Figure 1.4 : Principales étapes du pipeline bio-informatique FROGS pour l'analyse des données générées par la métagénomique ciblée

Figure 1.5 : Vue schématique de l'approche de clustering des séquences

Figure 1.6 : Processus écologiques contribuant à l'assemblage des communautés microbiennes des plantes

Figure 1.7 : Rapport entre les cellules et les gènes microbiens et humains dans le corps humain

Figure 1.8 : Arbre représentant la diversité génétique de 3 010 échantillons de *Oryza sativa*

Figure 1.9 : Structure des communautés microbiennes dans différents compartiments végétaux.

Figure 1.10 : Représentation d'une coupe transversale de racine de riz illustrant les différents compartiments hébergeant les communautés microbiennes

Figure 1.11 : Schéma d'une section de racine de riz présentant la structure de la rhizosphère, du rhizoplan et de l'endosphère dans un sol de rizière inondée

Figure 1.12 : Stade de développement du riz et dynamique de l'assemblage du microbiote racinaire du riz

Figure 1.13 : Schéma présentant les facteurs biotiques et abiotiques influençant la structure des communautés microbiennes des plantes

Figures du Chapitre II

Figure 2.1 : Localisation et présentation de l'agrosystème rizicole en terrasses de l'ethnie Hani de la vallée de Honghe

Figure 2.2 : Plan de la zone d'échantillonnage

Figures du Chapitre III

Figure 3.1 : 1) Station expérimentale du Mas d'Adrien : la parcelle « A » correspond au dispositif de l'année 2017 ; la parcelle « B » correspond au dispositif de l'année 2018 ; 2) Photo des microparcelles de la parcelle « B » au stade floraison en 2017

Figure 3.2 : Schéma des dispositifs expérimentaux de 2017 (1) et 2018 (2)

Figure 3.3 : Périodes d'échantillonnage sur les deux années de prélèvements

Figure 3.4 : Schéma des compartiments aérien et racinaire prélevés

Figure 3.5 : Schéma explicatif pour justifier la différence de traitement bio-informatique entre les échantillons de racines et les échantillons de la partie aérienne

Figure 3.6 : Représentation de la dynamique de prévalence virale d'oryza sativa alpha endornavirus

Figure 3.9: Structure des communautés bactériennes endophytes du riz entre les compartiments végétaux, Analyse en composantes principales (PCoA) utilisant l'indice de Jaccard entre les échantillons

Figure 3.10 : Abondance différentielle des taxa entre les compartiments

Figure 3.11 : Diagramme en violon de la richesse bactérienne endophyte des tiges (A ; B) et des racines (C ; D) entre les stades de développement de la plante pour les années de prélèvement 2017(A ; C) et 2018 (B ; D)

Figure 3.12 : Dynamique de structuration des communautés bactériennes endophytes des tiges et des racines de riz en fonction des stades de développement de la plante pour les années de prélèvement 2017 (A ; C) et 2018 (B ; D). Analyse en coordonnées principales (PCoA) sur l'indice UniFrac non pondéré

Figure 3.13 : Abondance différentielle des taxa endophytes des racines entre les stades de développement

Figure 3.14 : Abondance différentielle des taxa endophytes des tiges entre les stades de développement

Figure 3.15 : Dynamique de structuration des communautés bactériennes endophytes des tiges et des racines de riz en fonction des stades de développements de la plante, Analyse en composantes principales (PCoA) sur l'indice UniFrac non pondéré

Tableaux du Chapitre 2

Tableau 3.1 : Données disponibles sur l'estimation de la résistance à la pyriculariose en conditions de laboratoire

Tableau 3.2 : Analyse de la dynamique de prévalence virale l'oryza sativa alpha endornavirus pour les stades de développement floraison et maturation des grains en 2017 et tallage, floraison et maturation des grains en 2018 pour les six variétés de riz étudiées.

Tableau 3.3 : Analyse de la dynamique de prévalence virale d'oryza sativa alphaendornavirus pour les stades de développement floraison et maturation des grains en 2017 et tallage, floraison et maturation des grains en 2018 pour les six variétés de riz étudiées

Tableau 3.5 : Estimation de l'influence des différents facteurs sur la structure des communautés bactériennes endophytes du riz à l'aide d'un test de PERMANOVA basé sur l'indice de Jaccard (10 000 permutations)

Tableau 3.6 : Longueur du Cycle Semis-Epiaison (CSE) pour les six variétés suivies sur les deux années de prélèvements

Tableau 3.7 : Estimation de l'influence du stade de développement du riz sur la structure des communautés bactériennes endophytes des racines et des tiges de riz par une analyse de PERMANOVA basée sur l'indice de Jaccard et l'indice UniFrac non pondéré (10 000 permutations)

Tableau 3.8 : Estimation de l'influence de la variété de riz sur la structure des communautés bactériennes endophytes du riz sur chaque prélèvement réalisé par des analyses de variance multivariée PERMANOVA basées sur l'indice de Jaccard et l'indice UniFrac non pondéré (10 000 permutations)

Tableau 3.9 : Liste des OTUs fortement conservées (prévalence supérieure à 80%) et considérées comme appartenant au microbiote cœur des racines de riz pour chaque stade de développement et années de prélèvements

Liste des abréviations

ADN : Acide Desoxyribonucléique
ARN : Acide Ribonucléique
ARNr : Acide Ribonucléique ribosomique
ASV : Amplicon Sequence Variants
CFR : Centre Français du Riz
DALP : Direct amplification of length polymorphisms
DGGE : Denaturing Gel Gradient Electrophoresis
dsRNA : double strand RiboNucleic Acid
ITS : Internal transcribed spacer
NCBI : National Center for Biotechnology Information
FROGS : Find, Rapidly, OTUs with Galaxy Solution
GBS : Genotyping By Sequencing
HHRTS : Honghe Hani rice terraces system
ICTV : International Committee on Taxonomy of Viruses
OsEV : oryza sativa alphaendornavirus
OTU : Operationnal Taxonomic Units
pb : paires de bases
PCR : Polymerase Chain Reaction
PERMANOVA : Permutational multivariate analysis of variance
PFU : Plaque-Forming Unit
PGPR : Plant-Growth-Promoting-Bacteria
PNA : Peptide Nucléic Acid
qPCR : quantitative Polymerase Chain Reaction
rpm : rotation par minutes
RT-PCR : Reverse Transcriptase-Polymerase Chain Reaction
SSR : Single Sequence Repetition
ssRNA : single strand RiboNucleic Acid
SNP : Single nucleotide polymorphism
SRBSDV : Southern rice black-streaked dwarf virus
TAE : Tris-Acetate-EDTA
TGGE : Temperature Gel Gradient Electrophoresis
T-RFLP : Terminal Restriction Fragment Length Polymorphism
UV : UltraViolets
VANA : Virion-Associated Nucleic Acids

Chapitre 1 : Synthèse Bibliographique

L'étude des communautés microbiennes vise à identifier et quantifier les microorganismes présents à un instant donné et dans un habitat donné. Elle cherche aussi à comprendre l'évolution des structures de ces communautés dans l'espace et dans le temps en fonction de différents facteurs et à déterminer le rôle de chaque membre des communautés microbiennes au sein de l'hôte qu'il colonise.

1.1. Etude des communautés microbiennes par des méthodes moléculaires

En 1853, Joseph Leidy a publié le livre « A Flora and Fauna within Living Animals », qui est considéré comme l'origine des travaux de recherche sur le microbiote. Ont suivi ensuite jusqu'au début des années 1900 les travaux de Pasteur, Metchnikoff, Escherich, Kendall et autres qui ont permis de mieux comprendre les interactions hôtes-microorganismes (Aziz, 2009, Farré-Maduell & Casals-Pascual, 2019). Ces travaux ont bénéficié essentiellement de l'apport de techniques de culture des microorganismes qui ont permis, entre autres, de tester les postulats de Koch ou de mieux comprendre les mécanismes des interactions hôtes-microorganismes. Bien que ces approches culture-dépendante aient été largement renforcées grâce à l'amélioration des connaissances sur les conditions de culture des microorganismes, il reste une part importante de la diversité microbienne qui ne peut pas être décrite par la seule utilisation de ces approches. Les méthodes de culture standard permettent de décrire environ 1 % des bactéries de l'environnement. Cependant, par un effort de mise en culture et de multiplication des conditions de culture, il est estimé à ce jour qu'environ 50% des espèces d'une communauté bactérienne de l'intestin de mammifères ou de la rhizosphère d'*Arabidopsis thaliana* sont caractérisables par culture bactérienne (Lagkouvardos et al., 2017, Bai et al., 2015).

La mise en évidence dans les années 1970 par Karl Woese et collègues de plusieurs régions conservées au sein de la séquence du gène codant l'ARN ribosomique 16S (ARNr 16S) a bouleversé les connaissances sur l'identification et la diversité phylogénétique des bactéries (Woese & Fox, 1977, Fox et al., 1977). La découverte des propriétés d'identité et de variabilité des séquences de l'ARNr 16S a marqué le début de l'étude systématique des communautés microbiennes par des approches indépendantes de la culture. La classification des microorganismes reste aujourd'hui basée sur les principes d'alignements et de comparaisons des régions variables des séquences codant pour l'ARNr 16S (Figure 1.1).

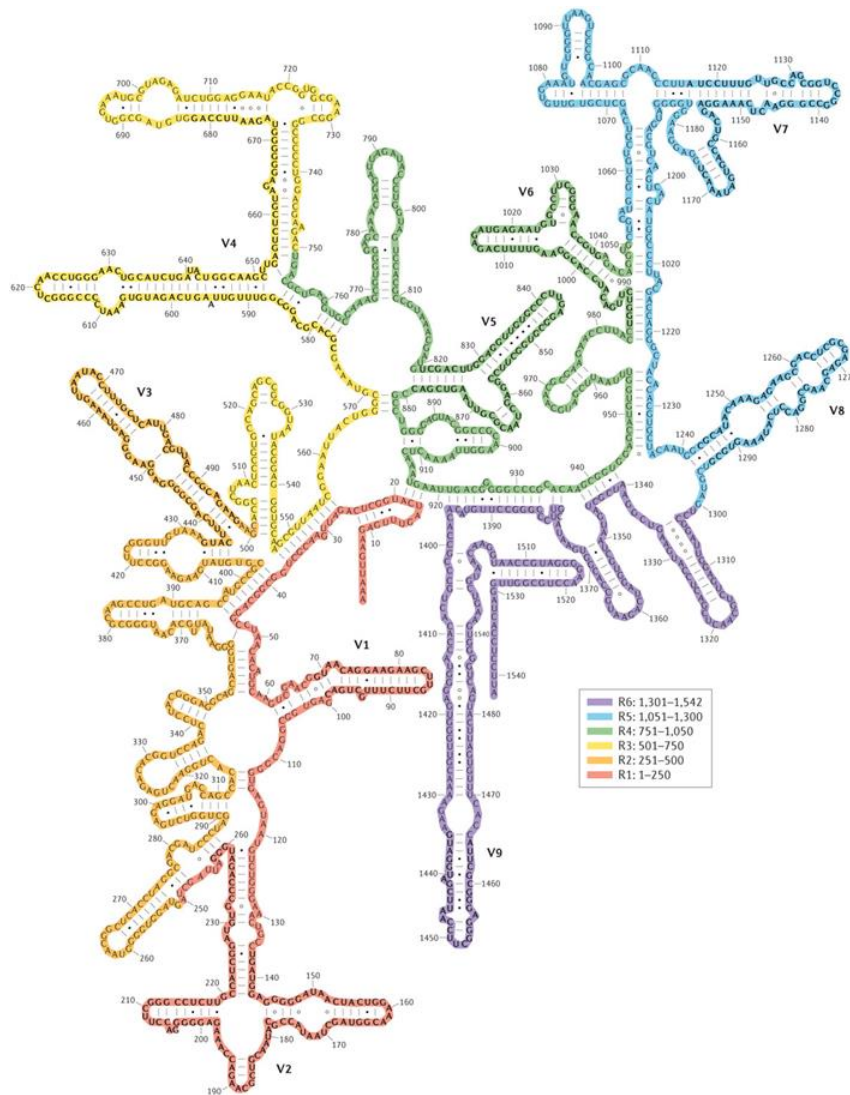


Figure 1.1 : Structure secondaire de l'ARNr 16S d'Escherichia coli (selon Yarza et coll. (Yarza et al., 2014)) présentant les différentes régions variables de cette molécule. En rouge, le fragment R1 comprenant les régions V1 et V2 ; en orange, le fragment R2 comprenant la région V3 ; en jaune, le fragment R3 comprenant la région V4 ; en vert, le fragment R4 comprenant les régions V5 et V6 ; en bleu, le fragment R5 comprenant les régions V7 et V8 ; et en violet, le fragment R6 comprenant la région V9

L'avènement de nouvelles approches moléculaires dans les années 1980, notamment la technique PCR (Polymerase Chain Reaction) et le séquençage Sanger ont contribué à l'essor de la compréhension des structures des communautés microbiennes. David Lane va être ainsi le premier à s'affranchir de l'obstacle de la culture pour l'identification des microorganismes en utilisant les techniques de PCR et de séquençage de l'ARNr 16S pour caractériser des mélanges bactériens sans avoir à les isoler (Lane et al., 1985, Yarza et al., 2014).

Handelsman et coll. proposeront par la suite, pour la première fois, le terme métagénomique pour désigner l'étude du contenu génétique de mélanges microbiens d'un échantillon

biologique par des méthodes moléculaires, sans passer par des étapes de culture et pouvant potentiellement donner accès aux génomes de tous les membres de la communauté (Handelsman et al., 1998). La métagénomique vise donc à caractériser l'ensemble des génomes d'une communauté microbienne d'un échantillon environnemental donné.

La métagénomique ciblée, pour l'étude des communautés microbiennes, est basée sur les travaux pionniers de Karl Woese et de David Lane et consiste en l'amplification PCR puis séquençage d'un gène orthologue que l'on suppose universel dans la communauté étudiée et suffisamment variable pour être considéré comme marqueur phylogénétique. L'ARNr 16S est un bon candidat pour l'étude des communautés bactérienne par métagénomique ciblée. Toutefois, avant l'essor des techniques de séquençage à haut débit, la métagénomique ciblée a tout d'abord fait appel à des techniques d'empreintes génétiques. Ces techniques permettaient de séparer sur gel les séquences des membres de la communauté microbienne en fonction de la composition de leurs séquences nucléotidiques et de comparer les profils représentant la structure des communautés. Ces approches incluent les électrophorèses sur gel en gradient dénaturant (Denaturing Gel Gradient Electrophoresis DGGE), en gradient de température (Temperature Gel Gradient Electrophoresis TGGE) et l'analyse du polymorphisme de longueur des fragments terminaux (Terminal Restriction Fragment Length Polymorphism T-RFLP) (van Elsas et al., 1998). Ces approches sont devenues rapidement obsolètes en raison de leurs faibles niveaux de résolutions, de leurs faibles capacités d'identification taxonomique des membres de la communauté et de l'impossibilité de multiplexer les échantillons limitant le nombre de répliques. L'arrivée des méthodes de séquençage à haut débit et la rapide diminution des coûts de séquençages à partir de 2008 ont permis de démocratiser l'étude des communautés microbiennes à différents domaines de recherche, dont l'étude des écosystèmes environnementaux.

L'ARNr 16S, bien que s'étant révélé un bon outil pour inventorier les communautés des procaryotes (bactéries et archée), présente cependant quelques limitations d'usage : un génome peut posséder plusieurs copies d'ARNr 16S (Louca et al., 2018, Case et al., 2007) ; les séquences des régions variables ne sont pas toujours assez discriminantes pour distinguer certaines espèces bactériennes entre elles (Johnson et al., 2019) et les assignations taxonomiques ne sont souvent réalisées qu'au niveau du genre. De nombreuses études ont montré l'intérêt d'utiliser d'autres marqueurs que le gène de l'ARNr 16S pour affiner l'assignation taxonomique des séquences. Ces marqueurs, tels que les gènes *rpoB* et *gyrB*, ont toujours une partie de leurs séquences hautement conservée permettant de créer des amorces

dites universelles pour amplifier le gène cible de la majorité des membres de la communauté analysée ; et une partie variable plus discriminante qui génère du polymorphisme de séquences. Bien que ces marqueurs alternatifs soient en copie unique dans le génome et qu'ils permettent une meilleure discrimination des membres de la communauté, l'ARN ribosomique 16S reste le marqueur phylogénétique le plus utilisé car il est fortement représenté dans les bases de données (SILVA, Greengene...) ce qui facilite l'assignation taxonomique des séquences.

A l'heure actuelle, l'analyse des communautés microbiennes utilise majoritairement : le gène de l'ARNr 16S pour les communautés bactériennes et les archées ; la région ITS (Internal Transcribed Spacer) pour les communautés fongiques.

1.1.1. Méthodologie de la métagénomique ciblée

La première étape de la métagénomique ciblée consiste au choix de la cible en fonction des communautés microbiennes analysées et de la technologie de séquençage.

1.1.1.1. Cibles génomiques pour analyser les communautés de procaryotes

Le gène de l'ARNr 16S qui est la cible de référence, pour les communautés bactériennes et les archées, a une longueur d'environ 1500 paires de bases (pb). Il est constitué de 9 régions variables d'une longueur allant de 30 à 200 pb (Yarza et al., 2014) (Figure 1.1). Ces régions variables V1 à V9 forment des boucles au niveau de la structure secondaire du gène de l'ARNr 16S (Figure 1.1). Elles sont séparées par 9 régions conservées composées de séquences complémentaires qui entraînent un repliement de la structure sur elle-même avec la formation de structures en « tiges » (Figure 1.1). La contrainte de repliement de la structure du gène de l'ARNr 16S, après sa réplication explique le faible taux de mutation des régions conservées par rapport aux régions variables (Yarza et al., 2014). De nombreux travaux ont ciblé différentes régions variables (Alcon-Giner et al., 2017, Yang et al., 2016, Rintala et al., 2017, Willis et al., 2019), une étude récente a identifié la région V3-V4 comme la plus adaptée et la plus discriminante avec la technologie de séquençage Illumina Miseq (Graspeuntner et al., 2018). Cette région, qui est amplifiée avec les amorces universelles 341f et 785r (Thijs et al., 2017), permet d'obtenir un fragment de 444 paires de bases.

1.1.1.2. Cibles génomiques pour analyser les communautés de champignons

Les gènes de l'ARN ribosomique codant pour les sous-unités 18S, 28S ainsi que les espaceurs internes transcrits (ITS) situés entre 2 sous-unités, sont le plus souvent utilisées pour cibler

l'amplification des microorganismes eucaryotes. Cependant, comme les cibles des gènes de l'ARN ribosomique 18S et 28S sont moins résolutive que les ITS1 et ITS2 (Iwen et al., 2002), les ITS sont majoritairement préférées pour la caractérisation des communautés fongiques (Figure 1.2). Des études portant sur les communautés fongiques ont été réalisées indépendamment sur les deux ITS, mais une récente qui conseille la région ITS2 comme étant la meilleure cible pour identifier les membres des communautés fongiques au niveau de l'espèce (Yang et al., 2018).

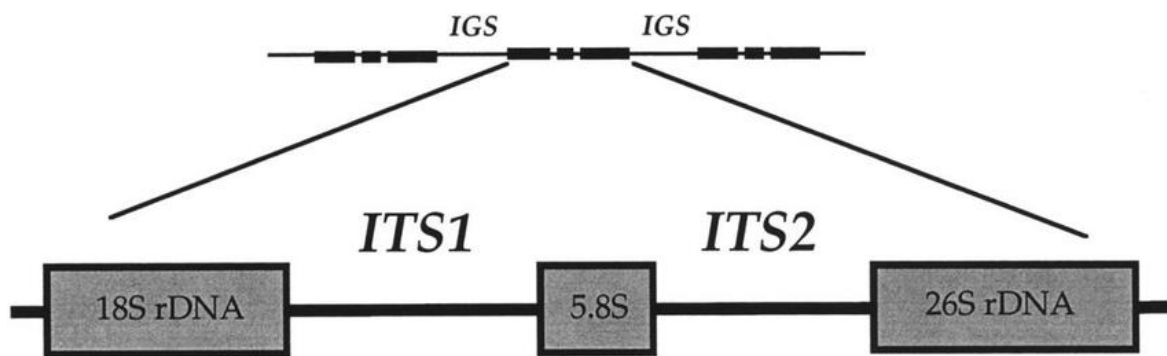


Figure 1.2 : Localisation génomique des espaces internes ITS1 et ITS2 (Shaw et al., 2002)

1.1.1.3. Technologies de séquençage à haut débit

Il existe à ce jour quatre technologies de séquençage adaptées aux approches de métagénomique ciblée, toutes proposées par la société Illumina. Les technologies Hiseq et Novaseq permettent le séquençage sens et anti-sens de fragments de 150 paires de bases et l'obtention respective de 2 et 10 milliards de séquences. Les technologies Miseq V3 et V4 permettent le séquençage sens et anti-sens, de fragments respectivement de 250 et 300 paires de bases et l'obtention de 15 et 25 millions de séquences. Bien que les technologies Miseq V3 et V4 ne permettent pas d'atteindre les profondeurs de séquençages des technologies Hiseq et Novaseq, elles sont plus adaptées aux études des communautés microbiennes ciblant les gènes de l'ARNr 16S et de la région ITS en raison des tailles d'amplicons couverts par ces technologies. Les produits d'amplification (amplicons) séquencés par les technologies Hiseq et Novaseq ne sont pas assez longs pour présenter suffisamment de variations de séquences et ainsi assurer une bonne assignation taxonomique des communautés microbiennes.

Après l'extraction de l'ADN total des échantillons (hôte et communauté microbienne), la séquence de la cible choisie (ARNr 16S ou ITS) est amplifiée par PCR avec les amorces universelles encadrant la région cible. Les amorces universelles possèdent à leur extrémité 5' une séquence d'environ 35 paires de bases appelée adaptateur. Une seconde amplification PCR permet l'hybridation des « barres codes » sur les adaptateurs de chaque échantillon rendant possible le multiplexage d'un très grand nombre d'échantillons. Tous les barres codes utilisés ont à leur extrémité 5' une séquence commune également appelée adaptateur qui permet l'hybridation des séquences barres codées sur la lame de séquençage et l'initiation de la réaction de séquençage. La lecture dans les deux sens de la cible permet d'obtenir une couverture totale de celle-ci avec un chevauchement d'environ 50 paires de bases permettant le rassemblement des deux sens de lecture de la séquence. Ainsi, les tailles maximales des séquences obtenues varient de 450 à 550 pb en fonction de la technologie Miseq utilisée.

1.1.2. La métagénomique globale ou métagénomique shotgun

La métagénomique globale est une approche à haut débit qui correspond au séquençage aléatoire du métagénome. Elle permet de caractériser les groupes taxonomiques des microorganismes présents dans l'échantillon ainsi que des fonctions associées à ces derniers car elle prend en compte l'ensemble du génome (Quince et al., 2017).

Malgré les avancées technologiques de séquençage permettant d'obtenir en une analyse plus de 10 milliards de séquences, l'analyse des communautés microbiennes associées à un hôte par métagénomique shotgun reste difficile à mettre en place. De plus pour une couverture de séquençage équivalente à la métagénomique ciblée, elle représente un coût plus important. La métagénomique shotgun s'intéresse à la totalité des organismes d'un échantillon qu'ils soient microscopique ou macroscopique. Lorsque l'étude de la communauté microbienne est réalisée chez un hôte eucaryote, l'ADN de son génome est ainsi également séquencé. Or celui-ci étant largement supérieur en concentration par rapport aux ADNs microbiens, il a tendance à dominer le résultat de séquençage et empêche une bonne couverture de la diversité microbienne.

Cependant, les technologies de séquençage offrant toujours plus de profondeur de séquençage, des projets de métagénomique shotgun ont récemment vu le jour pour caractériser les communautés microbiennes de différents écosystèmes. Ces projets restent cependant minoritaires par rapport à la métagénomique ciblée (Taur et al., 2018, Rampelli et al., 2019, Regalado et al., 2019).

En revanche, la métagénomique shotgun est particulièrement bien adaptée à l'étude des communautés virales (Roossinck et al., 2015). Les communautés virales ont pendant longtemps été délaissées des analyses du microbiote, qui se sont initialement focalisée sur les communautés bactériennes et dans un second temps sur les communautés fongiques. Cette absence d'intérêt pour les communautés virales était liée en partie à la controverse sur l'identité virale (Villarreal, 2009) : peuvent-ils être considérés comme des « êtres vivants » et être associés aux communautés du microbiote? Le statut biologique des virus reste encore flou (Forterre, 2016), mais leurs prises en compte dans l'étude des relations hôte/microbiote sont de plus en plus évidentes (Lima et al., 2019). Par ailleurs, les bactériophages (associés aux bactéries), les mycovirus (associés aux champignons) et potentiellement les phytovirus (associés aux plantes) ont un rôle dans la régulation des communautés microbiennes (Koskella & Brockhurst, 2014). Une seconde caractéristique des virus qui a freiné leurs considérations dans les projets de caractérisation des microbiotes est l'absence de gènes conservés entre toutes les entités virales (Roossinck et al., 2015). L'absence d'une région universellement conservée dans leurs génomes comme l'est l'ARNr 16S pour les bactéries, a rendu impossible leurs analyses par la méthode de métagénomique ciblée. En 2002, l'équipe de Forest Rowher a été la première à utiliser la métagénomique shotgun pour étudier le virome d'écosystèmes marins (Breitbart et al., 2002). Les résultats de cette étude et de toutes les études de métagénomique virale qui ont suivies ont bouleversé les connaissances sur la diversité virales avec la découverte d'une diversité extraordinaire d'entités virales qui n'avait encore jamais été observée jusqu'à ce jour par les techniques de culture ou de détection PCR (Zhang et al., 2018).

Toutefois, le rendement en séquences virales de ces approches de métagénomique shotgun a souvent été limité par le fait que les génomes viraux sont de petites tailles et ont des concentrations intracellulaires largement inférieures à celles des acides nucléiques de l'hôte et des autres microorganismes. Il a donc souvent été nécessaire de passer par une étape de semi-purification virale pour éliminer au maximum les acides nucléiques non-viraux avant le séquençage shotgun. Les tailles des particules virales bien inférieures à celles des cellules eucaryotes et procaryotes (Baker et al., 1999) permettent de les séparer des autres cellules contenues dans un échantillon par des méthodes de filtration et/ou de centrifugation (Thurber et al., 2009). Par ailleurs les ultrastructures des virus avec la présence de capsides protégeant le matériel génétique qu'elles contiennent permettent le traitement des échantillons biologiques par des nucléases afin d'éliminer les ADNs et les ARNs non viraux libres dans l'échantillon (Allander et al., 2001). Ces deux caractéristiques ont permis de réaliser des semi-purifications

de particules virales contenant la majorité des virus d'un échantillon (Roossinck et al., 2015). Certains virus de tailles supérieures à des bactéries n'ont cependant pas pu être étudiés après ces traitements (Raoult et al., 2004, Philippe et al., 2013). Le séquençage des semi-purifications virales qui contiennent peu de séquences nucléotidiques non virales a ensuite permis d'étudier le virome (ensemble des acides nucléiques viraux d'un échantillon) par des approches de métagénomique shotgun (Breitbart & Rohwer, 2005, Jones et al., 2005).

Mon équipe de travail a développé une approche de métagénomique shotgun permettant la description du virome associée à des espèces végétales (Filloux et al., 2015). Le principe de cette approche est de libérer les particules virales contenues dans la plante par un broyage de celle-ci puis d'éliminer les acides nucléiques de la plante hôte ou au moins en réduire la quantité. La purification des particules virales est réalisée en trois étapes i) des filtrations et centrifugations pour éliminer les cellules eucaryotes, ii) des ultracentrifugations pour concentrer les particules virales et iii) des digestions par des nucléases pour éliminer les ADNs et ARNs non viraux. Les ARNs et ADNs restants dans les purifications virales sont extraits conjointement (Hayes et al., 2017). Une rétrotranscription des ARNs viraux est réalisée à l'aide d'amorces aléatoires marquées par des barres codes. Cette étape est suivie par une réaction avec le fragment de Klenow puis par une amplification PCR visant les barres codes inclus dans les étapes précédentes. Un séquençage shotgun via la technologie Illumina Hiseq 2500 permet enfin l'obtention de nombreuses séquences virales d'environ 150 paires de bases qui seront ensuite assemblées par traitement bio-informatique.

1.1.3. Séquençage de longues séquences nucléotidiques

La troisième génération de technologie de séquençage « long reads » a cherché à s'affranchir i) des biais et des difficultés liées à l'amplification PCR, ii) de la longueur limitée des lectures de séquences qui induit des difficultés lors de l'assemblage ou de l'assignation taxonomique des séquences. Cette technologie permet le séquençage direct d'une molécule d'ADN dans son intégralité. Il s'agit des technologies PacBio ou NanoPore (Schneider & Dekker, 2012, Rhoads & Au, 2015). L'utilisation de ces technologies dans l'étude des communautés microbiennes permet une meilleure assignation taxonomique, car ces technologies ne sont pas limitées par la longueur des séquences (Wagner et al., 2016). Toutefois, le taux d'erreurs encore important de la technologie NanoPore ne lui permet pas de rivaliser à l'heure actuelle avec les méthodes Illumina utilisées pour les séquençages ciblés ou shotgun. L'usage de ces technologies dans le

cadre de l'étude des communautés microbiennes reste conditionné à leurs progrès technologiques futurs et ne sera pas développé davantage dans ce manuscrit.

1.2. Etude des communautés microbiennes par des méthodes de culture

Avant l'avènement des techniques de séquençage à haut débit, l'étude des flores microbiennes était presque exclusivement réalisée à l'aide de cultures, d'isolements et d'identifications morphologiques et biochimiques des bactéries et champignons. La culture des communautés microbiennes nécessite de maîtriser les conditions de culture spécifiques à chaque microorganisme pour permettre leurs multiplications et l'obtention d'un isolat en conditions de laboratoire. Les milieux complexes et les interactions entre microorganismes rendent difficile la mise en culture de l'ensemble des communautés microbiennes. Jusqu'en 2010, on estimait que seulement 20 % du microbiote intestinal humain était pris en compte par les méthodes de culture. Les 80% restants étaient considérés comme « non cultivables » ou non couverts par l'effort d'échantillonnage (Sekirot et al., 2010). La proportion élevée des microorganismes non cultivables peut être due à plusieurs facteurs tels qu'une dépendance à des conditions de culture très contraignantes (niveau d'oxygène, combinaison complexe de substrats et de microéléments, interactions avec un hôte ou d'autres microorganismes...). Il n'est pas suffisant de mimer les conditions de culture du milieu étudié, car certains microorganismes font partie de complexes microbiens impliquant des échanges de métabolites entre microorganismes pour permettre leurs croissances. Dans certaines conditions de culture, d'autres microorganismes peuvent avoir des effets antagonistes sur les autres membres de la communauté. La vitesse de croissance est un facteur important à prendre en compte car des microorganismes à croissance rapide peuvent masquer la présence de microorganismes à croissance plus lente ou en état de dormance et n'ont pas le temps de former des colonies de taille suffisante pour être détectées.

Récemment le développement de méthodes intensives de culture combinant différents milieux et différentes conditions de culture associées à une identification à haut débit par spectrométrie de masse MALDI-TOF rapide, précise et peu coûteuse à améliorer drastiquement les approches de culturomique (Lagier et al., 2018). Ces améliorations technologiques ont, en effet, permis l'identification dans l'intestin humain de nombreuses nouvelles bactéries qui n'avaient jusqu'alors jamais été observées par les techniques de culture bactérienne classiques ou même par les techniques de séquençage métagénomique dernière génération (Bilen et al., 2018, Lagier

et al., 2015, Lagier et al., 2018). Avec la « renaissance » de la culturomique, il a été montré la complémentarité des méthodes de culture et de métagénomique dans une caractérisation fine du microbiote. En effet, bien que les technologies de séquençage d'aujourd'hui permettent d'échantillonner des millions de fois la communauté microbienne pour en estimer sa diversité, les microorganismes avec les abondances les plus faibles ont généralement été moins bien couverts par les méthodes de métagénomique. Un faible niveau de recouvrement a ainsi été observé entre les deux méthodes (Lagier et al., 2012) (Figure 1.3) suggérant la nécessité de mener de front les deux approches pour caractériser au mieux le microbiote.

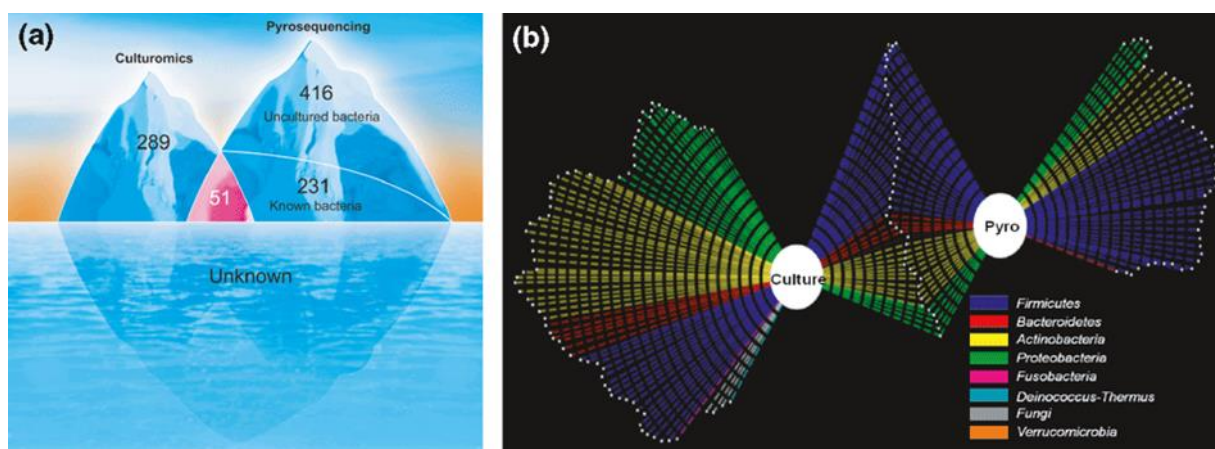


Figure 1.3 : **Illustration du recouvrement des résultats de la caractérisation des compositions des communautés microbiennes identifiées par une approche culture dépendante moderne et par une approche de pyroséquençage.** Ces travaux ont été réalisés par Lagier et coll., sur les bactéries dans l'intestin humain. (a) Les deux "icebergs" représentent les 340 espèces bactériennes cultivées et les 698 phylotypes identifiés par pyroséquençage. Le chevauchement entre les deux approches ne représente que 51 espèces, il est indiqué en rose. Sous le "niveau de la mer" est une estimation de la partie inconnue du microbiome de l'intestin humain. (b) Distribution taxonomique des microorganismes identifiés par une approche de culture et par une approche de pyroséquençage. Les lignes pointillées colorées indiquent les genres bactériens et fongiques identifiés par une seule approche (culture ou pyroséquençage), ou les genres identifiés par deux approches (Lagier et al., 2012).

Les études des communautés microbiennes par culturomique ne se sont pas limitées au domaine médical. Depuis quelques années, des travaux de caractérisations des communautés associées aux plantes ont été menés par des approches culture-dépendantes (Hacquard et al., 2015, Gouba et al., 2013, Priya et al., 2015, Xia et al., 2015, Long et al., 2010). Et se sont soldées par de bons recouvrements entre les méthodes culture-indépendantes et les méthodes culture-dépendantes.

L'approche de culture des microbiotes offre des perspectives pour approfondir le rôle joué par les différents membres des communautés microbiennes au sein de l'holobionte végétal. La constitution de collections de référence représentant différentes communautés microbiennes associées à un hôte ou à un environnement, permet de générer des microbiotes synthétiques qui peuvent être utilisés dans des expérimentations de laboratoire pour tester certaines hypothèses. Ce type d'essais basé sur l'utilisation de microbiotes synthétiques a déjà généré plusieurs résultats d'importance. Il a par exemple été montré un établissement préférentiel du microbiote sur différents compartiments de la plante en fonction de sa niche d'origine (Niu et al., 2017a, Bai et al., 2015) ou un effet sur le bon fonctionnement de la plante hôte (maïs) (Niu et al., 2017a).

1.3. Caractérisation de la diversité de l'hôte par typage génétique

L'étude des interactions hôte-microorganismes demande d'avoir une bonne connaissance de la composition des communautés microbiennes associées à l'hôte, mais également des connaissances sur la structure génétique de l'hôte. En effet, l'hôte peut impacter la structure des communautés microbiennes qui lui sont associées. Dans le cas des plantes, l'étude de la diversité des hôtes peut être réalisée par des approches de botanique, mais des approches utilisant des marqueurs moléculaires permettent une meilleure identification de la diversité génétique des hôtes.

Les premières études permettant le génotypage des individus étaient basées sur l'analyse des polymorphismes de longueur des fragments de restriction (Restriction Fragment Length Polymorphism, RFLP) (Bernatzky & Tanksley, 1986). Depuis l'utilisation de la technique RFLP pour le génotypage des plantes, de nombreuses autres techniques moléculaires ont été développées utilisant par exemple, les répétitions de séquences simples (SSR) (Litt & Luty, 1989), ou l'amplification directe du polymorphisme de longueur (Direct Amplification of Length Polymorphisms ou DALP) (Desmarais et al., 1998).

L'amélioration du débit de séquençage a permis l'utilisation des polymorphismes mononucléotidiques (Single Nucleotide Polymorphism, SNP) comme marqueurs pour le génotypage (Wang et al., 1998) non plus seulement sur des cibles spécifiques du génome, mais sur l'intégralité de celui-ci. Cette méthode de génotypage par séquençage appelé GBS (Genotyping By Sequencing) (Elshire et al., 2011) est réalisée après extraction de l'ADN de l'hôte et digestion de celui-ci par des enzymes de restriction. Les fragments d'ADN génomique

sont ensuite identifiés par ajout de barre codes pour être multiplexés et être séquencés par la technologie Illumina Hiseq qui permet un séquençage shotgun de séquences de 150 paires de bases. L'analyse bio-informatique des séquences obtenues par alignement sur un génome de référence permet d'identifier des milliers de marqueurs de diversité qui servent à générer des indices de diversité génétique des échantillons. La technique GBS a par ailleurs été couramment utilisée comme un outil de sélection assistée par marqueurs pour l'amélioration des plantes (He et al., 2014).

1.4. Analyses bio-informatiques

La quantité de séquences générées par les nouvelles méthodes de séquençage à haut débit produisant entre plusieurs millions et plusieurs milliards de séquences par analyse (run) a entraîné un développement important de méthodes informatiques pour gérer et analyser ces données produites en masse.

1.4.1. Métagénomique ciblée

Il existe plusieurs pipelines d'analyses bio-informatiques pour l'analyse des données générées par la métagénomique ciblée (FROGS (Escudié et al., 2017), QIIME (Caporaso et al., 2010), MOTHUR (Schloss et al., 2009)). Ces pipelines reposent tous sur les mêmes principes de tri, de nettoyage et de simplification des données issues du séquençage. La figure 1.4 présente les principales étapes de traitements bio-informatiques des données de séquençage réalisées avec le pipeline FROGS pour Find, Rapidly, OTUs with Galaxy Solution.

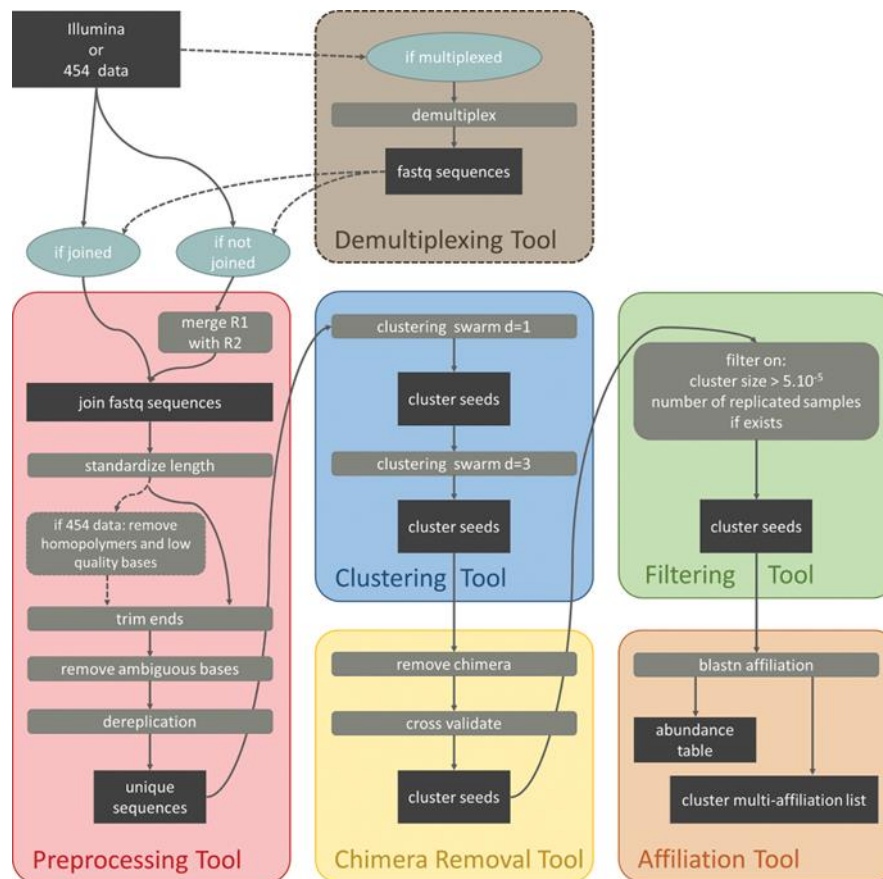


Figure 1.4 : Principales étapes du pipeline bio-informatique FROGS pour l'analyse des données générées par la métagénomique ciblée (Escudié et al., 2017).

1.4.1.1. Démultiplexage

Le démultiplexage a pour but de réattribuer toutes les séquences à leurs échantillons d'origine. Une recherche des barre codes est réalisée pour toutes les séquences afin obtenir des listes de séquences par échantillon.

1.4.1.2. Preprocessing

Le preprocessing a pour but de fusionner les lectures sens et antisens et d'éliminer les séquences de mauvaise qualité. Les séquences ne faisant pas la taille attendue sont supprimées : les séquences trop courtes ne peuvent pas permettre la fusion des séquences sens et antisens ; les lectures trop longues qui sont potentiellement des chimères sont aussi supprimées. Enfin les séquences ne permettant pas de retrouver les amorces d'amplification PCR ou celles contenant trop de bases dégénérées indiquent une mauvaise qualité de lecture sont aussi écartées de l'analyse. Enfin, une étape de déréplication permet de diminuer le nombre de séquences à manipuler en ne gardant qu'un seul exemplaire de toutes les séquences identiques, ce qui permet d'accélérer les calculs des étapes suivantes.

1.4.1.3. *Clustering*

Il existe plusieurs méthodes de clustering (aussi appelé débruitage - denoising). Le principe repose sur la réduction du nombre de séquences à traiter en corrigeant le bruit introduit dans les fragments séquencés ou en agglomérant les séquences proches. En effet, pour corriger différents biais techniques (erreurs de la polymérase lors de la PCR ou erreur de séquençage) qui génèrent de la diversité, le clustering des séquences permet de regrouper les séquences qui ont une forte identité pour ne conserver qu'une séquence représentative de tous les variants (souvent la plus abondante du cluster). Pendant longtemps, le seuil d'identité de 97% proposé par Stackebrandt et Goebel en 1994 (STACKEBRANDT & GOEBEL, 1994) permettait de différencier les différentes espèces microbiennes (Schloss & Handelsman, 2005, Seguritan & Rohwer, 2001, Westcott & Schloss, 2017). Ce seuil de 97% jugé comme arbitraire a été remplacé par plusieurs autres algorithmes de clustering (Edgar, 2018, Zou et al., 2018). La méthode que nous avons utilisée dans ce travail de thèse est appelée SWARM v2 et a été développée récemment par un chercheur de notre équipe de recherche (Mahe et al., 2015). La méthode d'agrégation des séquences développée dans l'algorithme SWARM v2 permet de regrouper des séquences, non pas à partir d'un seuil d'identité, mais en agrégeant progressivement les séquences qui ont jusqu'à une différence nucléotidique entre elles tant que leurs abondances sont décroissantes (Figure 1.5). Il est supposé par l'algorithme de calcul SWARM v2 que les séquences avec des erreurs de séquençages ne peuvent pas être plus abondantes que la séquence d'origine (Mahe et al., 2015). Les clusters formés sont appelés unités taxonomiques opérationnelles (OTUs) (Moyer et al., 1994). Seule la séquence la plus abondante des clusters est conservée. Elle est considérée comme représentative de la diversité de l'ensemble des séquences du cluster. Le paramétrage du clustering a une forte influence sur le nombre final d'OTUs et conditionne les indices de richesse des communautés microbiennes. Le clustering réduit fortement le nombre de séquences tout en conservant l'information d'abondance de celles-ci et rend possible l'assignation taxonomique qui était impossible à réaliser directement sur les séquences obtenues, en raison de leurs trop grands nombres.

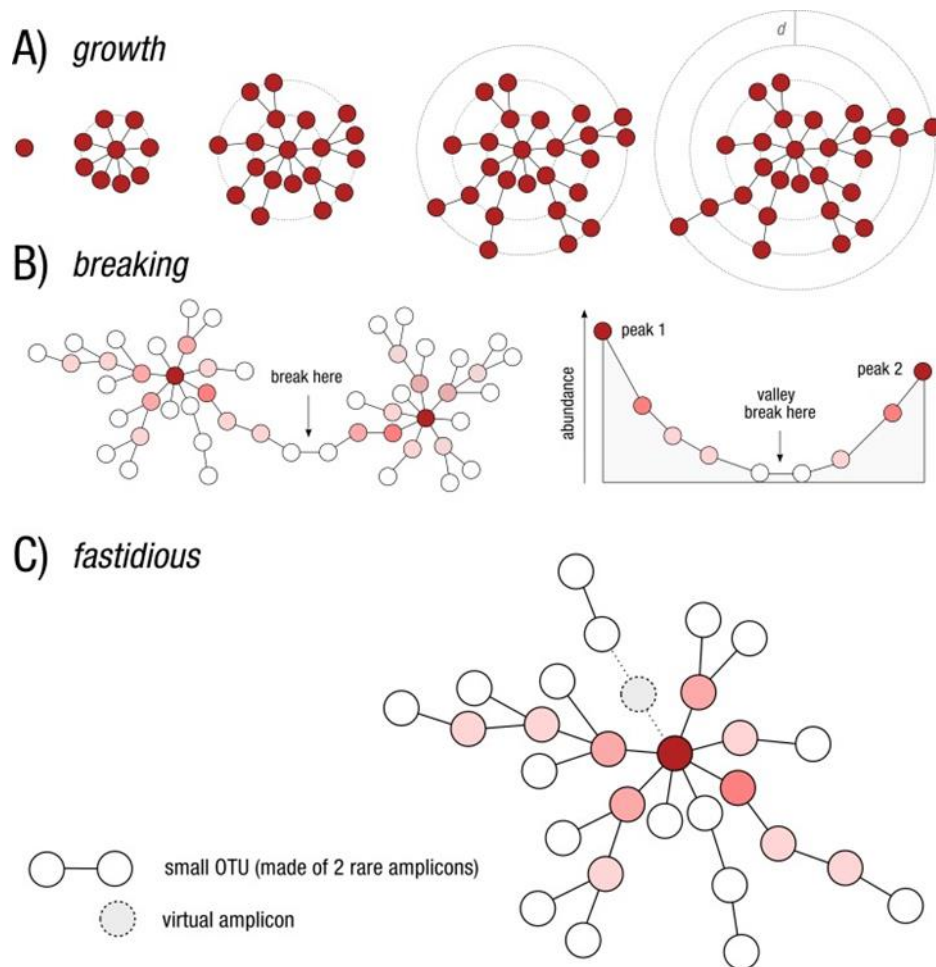


Figure 1.5 : Vue schématique de l'approche de clustering des séquences Swarm v2. (A) Swarm v2 regroupe les séquences de manière itérative en utilisant un petit seuil local « d » choisi par l'utilisateur, permettant aux OTUs de se développer jusqu'à leurs limites naturelles, où aucune autre séquence ne peut être ajoutée avec le seuil « d ». (B) Swarm v2 prend en compte l'abondance de chaque séquence. Plus le rouge est foncé, plus l'abondance est élevée. Le clustering des séquences en grappes est défini en ne permettant pas une augmentation de l'abondance dans le clustering des séquences. L'algorithme SWARM v2 estime que les erreurs de séquençage ne peuvent pas produire des séquences plus abondantes que leurs séquences d'origines. (C) L'option fastidieuse évite les sous-groupes (par exemple, la production de petites OTUs tels que les singletons et les doubletons) en postulant l'existence d'amplicons virtuels de liaison pour greffer les petites OTUs sur les plus grands (Mahe et al., 2015).

De nouvelles méthodes dédiées au débruitage des séquences ont émergé récemment sans réaliser de clustering des séquences. L'objectif est de corriger les erreurs de séquences dues aux divers biais méthodologiques plutôt que de les agréger entre elles (Callahan et al., 2017, Eren et al., 2013). Ces nouvelles méthodes permettent d'obtenir non pas des OTUs, mais des ASVs (Amplicon Sequence Variants) ou oligotypes. Bien que quelques articles scientifiques mettent en avant la supériorité de la méthode ASV pour discriminer les microorganismes, un récent article publié dans la revue Nature communications (Johnson et al., 2019) met en garde contre la conclusion selon laquelle la quantification effectuée par la méthode ASV est préférable aux approches plus traditionnelles basées sur les OTUs. En effet, il est démontré qu'étant donné que

de nombreuses bactéries contiennent de multiples copies variantes de l'ARNr 16S dans leur génome, les méthodes de clustering ASV ont tendance à surestimer la diversité des communautés bactériennes (Johnson et al., 2019). Suite à ces conclusions et à la large utilisation des méthodes de clustering en OTUs, nous avons fait le choix de ne pas utiliser la méthode ASV dans ce travail.

1.4.1.4. Déchimérisation

L'amplification PCR est connue pour introduire des erreurs ponctuelles de séquences, et pour produire des séquences chimériques (Wang & Wang, 1996). Les chimères sont des hybrides de deux (ou plus) séquences. Leurs formations sont liées à l'hybridation de deux séquences différentes au cours de l'étape d'élongation de la PCR. Les chimères sont séquencées comme tous les autres amplicons. Les chimères sont des sous-produits de PCR et ne constituent que des OTUs de faibles abondances, pouvant cependant représenter une part très importante du nombre d'OTUs (Bonder et al., 2012, Bachy et al., 2013). Les chimères peuvent être éliminées par deux méthodes : une méthode d'assemblage *de novo* et une méthode d'assemblage sur référence.

La déchimérisation *de novo* estime que les séquences qui ont permis la formation des chimères sont toujours plus abondantes que les séquences de la chimère formée. Les chimères sont détectées suite à un alignement de toutes les séquences d'un même échantillon. Les séquences identifiées comme mosaïques, c'est-à-dire qui contiennent des alignements partiels de plusieurs séquences, dont les séquences d'origine sont toutes plus abondantes, sont éliminées.

La déchimérisation sur référence estime que les séquences qui ont permis la formation des chimères sont présentes dans les bases de données. Un alignement sera réalisé comme pour la déchimérisation *de novo*, mais cette fois l'alignement de toutes les séquences des échantillons sera réalisé sur une base de données de référence (Silva, Greengenes...). Les séquences identifiées comme mosaïques, sont éliminées. La déchimérisation sur référence est moins utilisée, car les banques de références contiennent des chimères qui ne pourront donc pas être détectées par cette approche (Mysara et al., 2015).

1.4.1.5. Assignment taxonomique

Après toutes ces étapes de nettoyage et de décomplexification des données, les séquences OTUs sont annotées en recherchant des séquences similaires dans les bases de données de référence : Silva, Greengenes ou RDP pour les procaryotes et UNITE pour les champignons via des

recherches de similarités par l'algorithme BLAST (Altschul et al., 1990). Le lien phylogénétique entre les différentes OTUs peut être estimé en procédant à un alignement multiple puis en reconstruisant des arbres phylogénétiques par des approches phénétiques ou probabilistes. Finalement, une matrice regroupant l'abondance de chaque OTU pour chaque échantillon avec les résultats de l'assignation taxonomique est générée. Cette matrice appelée table d'OTUs sera utilisée pour les calculs des différents indices de diversités.

1.4.2. Métagénomique virale shotgun : analyses bio-informatiques des données de séquençages du virome par l'approche VANA

Les données de séquençages produites par la métagénomique VANA shotgun dans le cadre de la caractérisation du virome doivent subir plusieurs étapes de traitements bio-informatiques avant d'arriver à l'inventaire des virus présents dans l'échantillon. Les premières étapes du traitement des données de métagénomique shotgun sont identiques à la métagénomique ciblée.

1.4.2.1. Démultiplexage et preprocessing

Une étape de recherche des barre codes est réalisée pour démultiplexer les données avec l'algorithme Agrep (Wu & Manber, 1992). Une élimination des séquences estimées comme étant de mauvaise qualité est ensuite réalisée. Etant donné que toutes les séquences doivent avoir un double barre codes en 5' et en 3' spécifique à chaque échantillon, toutes séquences avec des mélanges de barre codes sont supprimés. Les barre codes et les amorces dégénérées utilisées pour l'amplification des séquences sont alors éliminés (étape de trimming) en utilisant l'algorithme Cutadapt (Martin, 2011).

1.4.2.2. Assemblage des séquences

Le séquençage shotgun produit un grand nombre de séquences de 150 paires de bases qui sont aléatoirement réparties sur tous les génomes viraux de l'échantillon. A ce stade, une assignation taxonomique des séquences serait longue et peu informative en raison du très grand nombre de séquences et de leurs faibles longueurs. Un assemblage *de novo* des séquences permet après alignement de regrouper les séquences se chevauchant provenant théoriquement du même génome en une seule séquence contigüe appelée contig. Cette étape réduit fortement le nombre de séquences en ne gardant qu'une séquence représentative de chaque contig formé. L'augmentation de la longueur des séquences après assemblage améliore la qualité de

l'assignation taxonomique. Dans certains cas l'assemblage peut permettre d'obtenir des génomes viraux entiers (Fancello et al., 2012).

1.4.2.3. Assignment taxonomique des séquences

Un alignement de toutes les séquences est réalisé sur un génome de référence de l'hôte, dans notre cas un génome d'*Oryza sativa japonica* pour éliminer toutes les séquences associées à la plante hôte. L'élimination des séquences de l'hôte qui sont non informatives améliore le temps de calcul pour l'assignation taxonomique des autres séquences.

Une recherche de similarité par alignement de toutes les séquences et contigs avec les séquences disponibles dans la base de données GenBank est réalisée avec les algorithmes BlastN et BlastX. Les contigs et les séquences simples (singleton) attribuées à des séquences virales connues dans GenBank sont alors sélectionnés et regroupés dans une table de données. Les séquences associées à des acides nucléiques eucaryotes et procaryotes sont supprimées.

1.4.3. Analyses bio-informatiques des données de géotypage des plantes par GBS

1.4.3.1. Démultiplexage

L'étape de démultiplexage des séquences est identique à celles détaillées dans les précédentes méthodes (décrite dans la présentation de la métagénomique ciblée et shotgun).

1.4.3.2. Alignement sur un génome de référence

Un alignement de toutes les séquences sur un génome de référence, si possible de la même espèce que les échantillons analysés, permet d'identifier des zones du génome présentant du polymorphisme de séquences de type SNP, également marqueur de diversité. Le séquençage shotgun fournit un grand nombre de séquences distribuées aléatoirement sur le génome de la plante hôte. Le génome du riz *Oryza sativa indica* avec près de 390 millions de paires de bases (Du et al., 2017) ne peut pas être couvert dans son intégralité par tous les échantillons. Seuls les marqueurs présentant moins de 20% de données manquantes sont conservés pour les analyses de diversité de la plante. Autrement dit sont conservés les marqueurs pour lesquels 80% des échantillons ont fourni une séquence. Une séquence représentative constituée de tous les marqueurs de diversité placés les uns à côté des autres (séquence concaténée) est utilisée pour les alignements et les analyses de diversité génétique des échantillons.

1.5. Limites des méthodes d'analyses des communautés microbiennes

Les méthodes basées sur le séquençage sont aujourd'hui massivement utilisées dans l'étude des communautés microbiennes. Elles proposent une solution rapide avec un coût modéré pour estimer le nombre d'espèces et leurs abondances relatives au sein d'un échantillon. Ces approches montrent parfois leurs limites (Quince et al., 2017). En effet, toutes les méthodes ciblant l'ADN ou l'ARN environnemental sont basées sur une extraction des acides nucléiques, et d'un séquençage d'ADN qui est ciblé après amplification PCR d'une séquence spécifique et sont toutes soumises aux mêmes biais : l'extraction d'ADN n'est pas exhaustive et l'efficacité d'extraction est variable suivant les échantillons potentiellement influant sur l'estimation de la diversité (Knudsen et al., 2016) ; la profondeur de séquençage et la qualité des bases de références génomiques pour l'écosystème étudié influencent également l'estimation de la diversité. Ainsi, un effort de séquençage insuffisant ne permet pas d'estimer la diversité des microorganismes les moins abondants. Malgré des améliorations continues des bases génomiques associées aux microorganismes, une grande partie de la diversité microbienne est par ailleurs à ce jour mal connue. Dénommée usuellement la matière noire (« dark matter »), elle correspond à toute la diversité de séquences nucléotidiques attribuées à aucun microorganisme connu (Rinke et al., 2013). Cette matière noire est plus importante dans l'estimation de la diversité des communautés virales, en raison de la plus grande complexité de diversité de séquences chez les virus et de la relative faiblesse des bases de données internationales (Krishnamurthy & Wang, 2017).

Enfin, des phénomènes de contaminations croisées ou « index hopping » liées à la technologie de séquençage Illumina sont connus (Valk et al., 2019, Nelson et al., 2014). Leurs survenues restent peu élevées. Cependant si les communautés étudiées présentent une faible diversité avec des populations très dominantes, le niveau d'index hopping peut s'avérer substantiel dans un grand nombre d'échantillons (Valk et al., 2019, Nelson et al., 2014).

1.5.1. Limites spécifiques à la métagénomique ciblée

En plus des limitations communes à toutes les méthodes de séquençage, l'approche de métagénomique ciblée est soumise aux biais de PCR et de choix de cibles. Le choix des cibles influence fortement l'estimation de la diversité microbienne :

- Les cibles choisies peuvent être plus ou moins discriminantes des différents membres de la communauté microbienne en fonction de la variabilité de la zone choisie et de la longueur de la séquence amplifiée (Rintala et al., 2017). Ces cibles ne sont pas toujours suffisantes pour obtenir une résolution taxonomique au niveau de l'espèce (plusieurs espèces peuvent avoir des séquences cibles identiques).
- Le choix des amorces dites universelles pour amplifier la cible et la réaction PCR peuvent biaiser l'estimation de la diversité microbienne. Les amorces sont choisies pour amplifier un maximum de taxa, mais elles peuvent montrer une certaine spécificité pour certains microorganismes qui seront préférentiellement amplifiés par rapport à d'autres. Certains microorganismes peuvent même ne pas être amplifiés (Aird et al., 2011, Bellemain et al., 2010).
- Les cibles les plus utilisées (16S/18S/ITS) ont un nombre de copies qui peut varier en fonction des espèces. De plus, les différentes copies de la cible peuvent présenter une hétérogénéité de séquences ce qui entraîne une surestimation de la diversité microbienne (Vetrovsky & Baldrian, 2013, Angly et al., 2014).
- L'amplification PCR et le séquençage peuvent être à l'origine de séquences non présentes dans l'échantillon de départ car les enzymes utilisées pour le séquençage peuvent être plus ou moins fidèles (Gury et al., 2008). Une autre source d'erreurs liée à l'amplification PCR est la formation de séquences chimériques (Bonder et al., 2012, Bachy et al., 2013). Le séquençage est aussi générateur d'erreurs (Nelson et al., 2014) qui peuvent être stochastiques ou systématiques. Elles sont difficiles à prendre en compte, mais leur importance est réduite par l'étape de clustering des séquences.

Une autre limitation de la métagénomique ciblée est qu'elle n'est pas quantitative. Il n'y a pas de lien direct entre la biomasse de l'individu et le nombre de séquences obtenues. Il est impossible d'avoir une approche quantitative en raison des multiples biais techniques présentés ci-dessus (nombre de copies variables de la cible, affinité des amorces variables entre les microorganismes, efficacité d'amplification variable...). Une approche semi-quantitative pour des échantillons homogènes est quant à elle envisageable. Il est possible d'estimer les différences d'abondances d'un taxon entre des échantillons. Cependant il n'est pas possible d'estimer les différences d'abondance entre les taxa d'un même échantillon (AMEND et al., 2010).

Enfin, la métagénomique ciblée ne donne accès qu'à un inventaire des ADNs microbiens contenus dans l'échantillon. Elle ne donne pas d'information sur les fonctions des communautés

microbiennes ni sur leur activité. Par exemple, les acides nucléiques non dégradés des microorganismes morts peuvent être pris en compte dans les analyses de diversité (Emerson et al., 2017). Ainsi, le rôle fonctionnel des communautés sera mieux abordé par les approches de métagénomique globale, ou métatranscriptomique ou par culture microbienne.

1.5.2. Limites de la métagénomique globale

La principale limite à la métagénomique globale pour l'étude des communautés bactériennes et fongiques est liée à la profondeur de séquençage pour couvrir la diversité de tous les génomes microbiens. Cette limite est d'autant plus importante pour les échantillons contenant les cellules de l'hôte e.g. le microbiote de l'endosphère de la plante, en raison de la fraction extrêmement élevée de l'ADN de l'hôte par rapport à l'ADN microbien (Pereira-Marques et al., 2019).

La métagénomique globale demande un assemblage des données de séquences qui est difficile et peut générer de nombreuses erreurs en raison de l'absence de génomes de référence pour tous les microorganismes. De plus, les nombreux génomes microbiens disponibles dans les banques de données génétiques ne sont souvent pas représentatifs de la communauté étudiée et correspondent souvent à des microorganismes modèles.

Cette observation est encore plus vraie pour l'étude des communautés virales en raison d'une faible disponibilité en génomes viraux. Le taux élevé de recombinaison pose également des difficultés pour l'assemblage *de novo* des séquences. Par ailleurs, la semi-purification virale VANA qui est une technique d'enrichissement des particules virales par filtration et centrifugations, peut biaiser l'estimation de la diversité virale. Il existe par exemple des virus d'une taille comparable à celle de bactéries (Halary et al., 2016) qui ne sont pas retenus par cette méthode.

L'approche VANA n'est pas une analyse de métagénomique shotgun classique. Pour pallier à la faible concentration en acides nucléiques des semi-purifications virales, une amplification sans a priori est réalisée avec des amorces dégénérées (Bernardo et al., 2018). Comme toutes les techniques d'amplification des acides nucléiques, l'amplification peut induire des biais liés à des erreurs de la polymérase, à la création de séquences chimériques etc. (Kim & Bae, 2011).

1.5.3. Limites des méthodes de culturomique

Les méthodes de culture ont permis d'identifier des microorganismes de faible abondance relative, souvent encore jamais observés par les méthodes de métagénomique (Lagier et al.,

2016, Lagier et al., 2012). A contrario, les travaux de Lagier et coll. ont également montré que les méthodes indépendantes de la culture permettaient d'accéder à une partie de la diversité bactérienne qui n'était pas couverte par les méthodes de culture en raison des difficultés à cultiver ces organismes (Lagier et al., 2012). Ces travaux ont démontré la complémentarité des deux approches dans la caractérisation des communautés microbiennes. Les avancées dans le domaine de la culture bactérienne et l'utilisation de la spectrométrie de masse MALDI-TOF, ont permis une augmentation du débit d'analyses et une diminution des coûts. La technologie MALDI-TOF a permis de réduire considérablement le temps nécessaire à l'identification bactérienne en réalisant des identifications directement à partir de colonies par des analyses de leurs profils peptidiques. Cette technologie est plus précise et plus rapide qu'une analyse de génotypage par séquençage type Sanger, mais demande de disposer d'une base de données de références (Tandina et al., 2016, Samb-Ba et al., 2014). Or, aujourd'hui il n'existe pas de bases publiques pour les spectres MALDI-TOF notamment pour les microbes environnementaux comme pour les bases génomiques (SILVA, Greengene...) et le travail de création de sa propre base de données est un travail long et fastidieux.

1.6. Estimation de la diversité des communautés microbiennes

Un des premiers objectifs des analyses des communautés microbiennes est de décrire leur diversité et leur composition. La diversité des communautés microbiennes peut être estimée à plusieurs niveaux (Xia et al., 2018, Whittaker, 1960, Ganeshaiah et al., 2019):

- La diversité locale α , représente la diversité à l'intérieur d'un système délimité (une niche, un échantillon). Elle permet d'estimer le nombre d'espèces ou taxa présents dans l'échantillon analysé.
- La diversité β est la variation de composition entre des assemblages locaux ou entre des assemblages régionaux. Elle a été définie originellement comme une mesure de changement de diversité le long d'un gradient environnemental. Elle permet de savoir quels sont les taxa partagés entre communautés ou de connaître les taxa uniques à un assemblage.
- La diversité γ est la diversité totale et représente la diversité dans un ensemble d'échantillons. Elle est partitionnée en deux composantes la diversité α et la diversité β .

La diversité observée est le résultat de différents processus qui s'exercent lors de l'assemblage d'un microbiote et vont générer de la variabilité spatio-temporelle entre ces communautés.

Vellend en 2010 a réalisé une synthèse conceptuelle en écologie des communautés en proposant quatre catégories de processus contribuant à l'assemblage des communautés (Vellend, 2010). Ces concepts qui sont présentés ci-dessous (Figure 1.6) ont été discutés par la suite dans le cadre de l'assemblage des communautés microbiennes (Zhou & Ning, 2017, Nemergut et al., 2013, Cordovez et al., 2019).

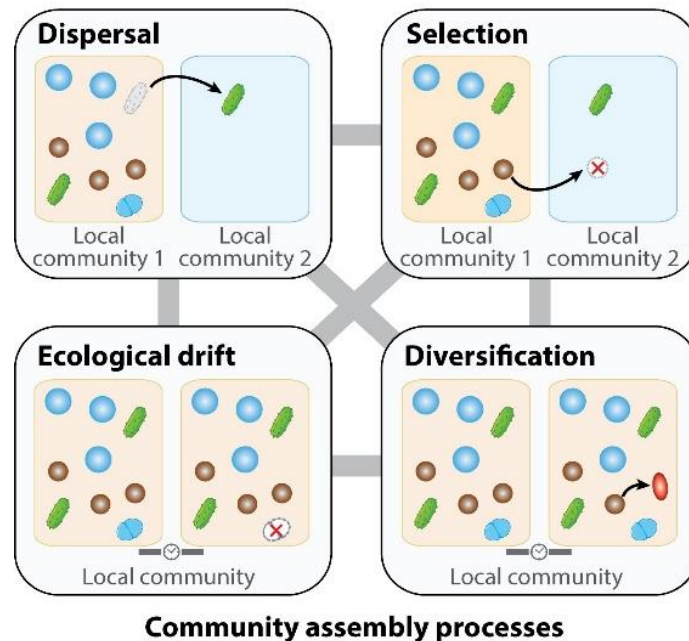


Figure 1.6 : Processus écologiques contribuant à l'assemblage des communautés microbiennes des plantes. Quatre processus éco-évolutifs structurent l'assemblage de la communauté : la dispersion ; la sélection ; la dérive écologique et la diversification. (Cordovez et al., 2019).

- La sélection

La sélection décrit des différences déterministes de « fitness » (terme anglo-saxon qu'on peut imparfaitement traduire en français par « valeur sélective » (Bolduc, 2012)) entre individus de différentes espèces. Elle apparaît quand les fréquences d'individus au sein d'une communauté varient entre elles et quand différents variants se reproduisent ou se multiplient à des taux différents. La sélection peut varier dans le temps et dans l'espace. En général, la coexistence d'espèces dépend d'un certain équilibre, avec des espèces différentes qui ont un avantage de fitness sous différentes conditions parfois spécifiées par des combinaisons de facteurs abiotiques environnementaux ou par les densités des espèces elles-mêmes (Vellend, 2010). En revanche, une part de cette diversité microbienne ainsi que les nombreuses interactions existant entre microorganismes (compétition, mutualisme, commensalisme) qui vont déterminer la structure des communautés microbiennes sont encore mal connues (Nemergut et al., 2013).

- La dérive écologique

C'est un processus stochastique important qui va générer des changements de composition et de diversité des espèces. La naissance, la mort, la production d'une descendance étant des processus stochastiques, les changements dans toute communauté finie auront aussi une composante stochastique. Une diminution de la taille effective d'une communauté liée à une perturbation de l'écosystème peut augmenter l'importance de la dérive. Cette situation est probablement rencontrée dans le cas des environnements de la phyllosphère et de l'endosphère qui présentent des densités et des niveaux de diversités des microorganismes connus pour être plus faibles qu'au sein de la rhizosphère (Cordovez et al., 2019). Par ailleurs, dans le cas d'une faible pression de sélection et d'une petite taille de communauté, les effets de la dérive peuvent surpasser ceux de la sélection.

- La spéciation ou diversification

La diversification est un processus évolutif qui permet de générer de nouveaux variants et qui est en équilibre avec la spéciation et l'extinction. Elle joue un rôle important dans la constitution du pool d'espèces ou de variants à l'échelle régionale qui en retour peut modifier les profils de communautés au niveau local suite à de la dispersion. A une échelle locale, la création d'une nouvelle espèce proprement dite est négligeable par rapport aux autres processus. Toutefois, la diversification joue dans certains cas un rôle crucial dans l'adaptation des microorganismes à un niveau très local. A titre d'exemple, les travaux menés sur un champignon inféodé au riz, *Harpophora oryzae*, ont montré comment une diversification génomique de ce microorganisme impliquant différents mécanismes génétiques (gains et pertes de fonctions, duplication de gènes, modifications des protéines impliquées dans les interactions hôte-microorganisme, mouvements d'éléments transposables, etc.) s'est traduite par la conversion d'un pathogène en un endophyte mutualiste du riz (Xu et al., 2014).

- La dispersion

La dispersion regroupe les processus de mouvement et d'établissement réussi d'organismes dans l'espace. Elle peut se faire par des modes actifs ou passifs et peut dépendre de facteurs déterministes et stochastiques (Zhou & Ning, 2017). Les processus de dispersion des microbes restent peu étudiés et les capacités de dispersion du microbiote restent aussi très discutées. Mais la dispersion est un facteur important influençant le pool régional d'espèces et la structure des

communautés associées. Un fort taux de dispersion peut homogénéiser la structure des communautés et réduire la variation entre elles. La dispersion peut aussi avoir un rôle très important dans l'assemblage du microbiote des plantes et par extension un effet fort sur le fonctionnement de l'holobionte végétal. A titre d'exemple, une étude récente focalisant sur le rôle du microbiote racinaire dans la protection du maïs contre le champignon pathogène *Fusarium verticillioides* a identifié l'importance d'un groupe restreint de sept souches bactériennes pour inhiber cet agent pathogène. Mais, cette étude est surtout considérée comme pionnière pour avoir montré l'importance de l'ordre d'introduction de ces sept souches bactériennes à la surface de la plante pour obtenir un effet maximum de contrôle du champignon pathogène, révélant ainsi l'importance des processus de succession et de dispersion (Mendes et al., 2013).

1.6.1. Indices de diversité

En se basant sur les tables d'abondances, plusieurs mesures de diversité peuvent être estimées. Dans les études du microbiote, les mesures de diversité intra- et inter-communautés sont communément réalisées. Il existe de nombreux indices de diversité pour décrire la diversité α et β d'un écosystème (Wagner et al., 2018). Dans ce paragraphe nous ne présenterons que quelques-uns d'entre eux qui ont été par la suite utilisés dans les travaux de cette thèse.

1.6.1.1. *Indices de diversité α*

Richesse spécifique

La richesse spécifique correspond à l'indicateur de diversité α le plus simple. Il mesure simplement le nombre d'entités différentes par échantillon ou par condition. Dans le cas du microbiote analysé par des approches de métagénomique ciblée, ce sont des taxa ou des OTUs. Cette mesure donne de l'importance aux espèces rares car toutes les espèces ont le même poids, quelles que soient leurs fréquences.

Indice de Shannon-Wiener

L'indice de Shannon a été défini à la fois par Claude Shannon et par Norbert Wiener en 1948. La diversité de Shannon prend en compte non seulement le nombre d'entités, mais également leur distribution, i.e. la proportion ou l'abondance de chaque entité (aussi appelé « équitabilité »). Un autre indice de diversité α appelé indice de Simpson accorde plus d'importance aux entités les plus abondantes.

L'indice de Shannon-Wiener le plus couramment utilisé est donné par la formule suivante (Spellerberg & Fedor, 2003) :

$$\text{Indice de shannon } H' = \sum_{i=1}^S P_i \ln P_i$$

Pi représente l'abondance relative du taxon

S = nombre total d'espèces

L'indice de Shannon permet d'exprimer la diversité en prenant en compte le nombre d'entités microbiennes et l'abondance au sein de chacune des entités microbiennes. Ainsi, une communauté dominée par une seule espèce aura un indice H moins élevé qu'une communauté dont toutes les espèces sont équitablement distribuées. La valeur de l'indice varie de 0 à $\log S$. La valeur 0 signifie qu'il y a une seule espèce ; une valeur proche de 0 signifie qu'une espèce domine très largement toutes les autres ; $\log S$ signifie que toutes les espèces ont même abondance. L'indice de Shannon est classiquement compris entre 0 et 5 pour les études de communautés microbiennes.

1.6.1.2. Indices de diversité β

De nombreuses mesures de distances de similarité ou de dissimilarité permettent de comparer les échantillons deux à deux, et de leur attribuer une valeur illustrant leurs ressemblances globales (Schroeder & Jenkins, 2018). Dans les analyses de diversité microbienne on distingue deux catégories d'indices de diversité β . Les coefficients de similarité binaire basés sur des données de présence/absence (Jaccard, UniFrac non pondéré) et des coefficients de similarité quantitative quand des mesures d'abondance relative sont disponibles (Bray-Curtis, UniFrac pondéré) (Xia et al., 2018). Par ailleurs, on distingue aussi des distances de similarité (ou dissimilarité) basées sur le nombre d'OTUs communes entre deux échantillons (Jaccard, Bray-Curtis) et d'autres qui sont basées sur la distance de similarité phylogénique des communautés microbiennes (UniFrac).

Distance de dissimilarité de Jaccard et de Bray-Curtis

Ces deux indices sont basés sur la différence de composition des communautés microbiennes et tiennent compte du taux de renouvellement des OTUs entre les échantillons.

La distance de Jaccard calcule le nombre d'OTUs communes entre deux échantillons. Ce calcul est réalisé entre toutes les paires d'échantillons et permet de créer une matrice de distances de similarité entre tous les échantillons. Cet indice est souvent transformé en distance de dissimilarité pour comparer les résultats avec la distance de dissimilarité de Bray-Curtis (Jaccard, 1912).

Le principe de la distance de dissimilarité de Bray-Curtis est le même que la distance de Jaccard à la différence qu'il tient compte de l'abondance des OTUs dans la comparaison (Bray & Curtis, 1957).

$$\text{Distance de Jaccard, } J = \frac{C}{C + A + B}$$

A = nombre d'OTUs uniquement présentes dans un échantillon donné (x) par rapport à un autre échantillon (y)

B = nombre d'OTUs uniquement présentes dans l'échantillon (y) par rapport à l'échantillon (x)

C = somme des OTUs entre les deux échantillons (x) et (y)

$$\text{Distance de dissimilarité de Bray - Curtis, } BC = 1 - \frac{2W}{D + E}$$

D = somme des abondances de toutes les OTUs dans un échantillon donné

E = somme des abondances des OTUs d'un autre échantillon

W = somme des valeurs d'abondance les plus faibles pour chaque OTU commune aux deux échantillons

Les distances de dissimilarité de Bray-Curtis et de Jaccard varient entre 0 (valeurs identiques pour toutes les espèces) et 1 (aucune espèce en commun).

Distance UniFrac pondérée et non pondérée

Les méthodes de calcul des distances UniFrac intègrent l'information des liens de parenté entre les membres des communautés microbiennes. Ces liens sont estimés par le calcul de la distance phylogénétique entre les membres des communautés microbiennes (Lozupone & Knight, 2005). Les distances UniFrac utilisent un arbre phylogénétique calculé après alignement des séquences de toutes les OTUs détectées. Le calcul de la distance UniFrac se base sur les longueurs des branches de l'arbre partagées entre deux échantillons. Il peut être pondéré ou non par les valeurs d'abondances des OTUs.

La distance UniFrac non-pondérée (unweighted UniFrac) ne tient pas compte de l'abondance des OTUs et ne considère que la présence/absence des OTUs. Elle est donnée par la formule suivante :

$$d_{UniFrac} = \frac{l_{unique}}{L_{totale}}$$

L_{unique} = somme des longueurs des branches propres à l'échantillon (x) par rapport à l'échantillon (y)

L_{totale} = somme de toutes les longueurs des branches de l'échantillon (x)

Dans le cas où deux échantillons partagent toutes leurs OTUs, $unique = 0$, donc $d_{UniFrac} = 0$. À l'autre extrême, lorsque deux échantillons ne partagent aucune OTU, $unique = totale = 1$, donc $d_{UniFrac} = 1$.

La distance UniFrac pondérée (weighted UniFrac) tient compte de l'abondance des OTUs dans l'estimation des différences, elle accorde plus d'importance aux OTUs abondantes (Lozupone et al., 2007). Elle est donnée par la formule suivante :

$$d_{WUniFrac} = \sum_i^n b_i \times \left| \frac{X_i}{X_T} - \frac{Y_i}{Y_T} \right|$$

n = nombre total des branches de l'arbre

b_i = longueurs de la branche i

X_i = nombre de séquences descendant de la branche i pour l'échantillon (x)

Y_i = nombre de séquences descendant de la branche i pour l'échantillon (y)

X_T = nombre total de séquences de l'échantillon (x)

Y_T = nombre total de séquences de l'échantillon (y)

1.7. De la notion de symbiose au concept de l'holobionte

Selon la première définition donnée par Anton de Bary en 1879, la symbiose du grec « symbiôsis » signifie vivre ensemble et qualifie les associations intimes et durables entre deux ou plusieurs espèces (Oulhen et al., 2016). La symbiose peut se traduire par l'association d'un hôte eucaryote à un microorganisme telle que retrouvée entre une plante et une bactérie (Martin et al., 2017) ou par l'association entre deux microorganismes, comme par exemple entre des bactéries endosymbiotiques et des cellules fongiques dans lesquelles elles se développent (Bastias et al., 2019). On distingue différentes relations de symbiose en fonction des coûts/bénéfices engendrés par les associations (Bass et al., 2019, Tshikantwa et al., 2018). Ainsi, on distingue : des symbioses mutualistes qui sont bénéfiques aux deux parties ; des symbioses commensales qui sont bénéfiques à une partie et sans effet pour le second acteur et enfin des symbioses parasitaires qui entraînent un coût pour une des deux parties. Les relations de symbioses d'un hôte eucaryote avec un ou des microorganismes qu'ils soient mutualistes, commensaux ou parasitaires peuvent être également qualifiées de facultatives lorsqu'elles ne sont pas nécessaires à la survie à l'une des deux parties ou obligatoires lorsque la survie de l'une des deux parties en dépend.

Suite à l'avènement des techniques de séquençage à haut débit, aux progrès des outils de bio-informatique et à la constitution de bases de données internationales et publiques de séquences universelles de gènes microbiens, le concept de symbiose a quelque peu évolué en prenant en compte non plus des duos ou des trios de partenaires symbiotiques (Van der Heijden et al.,

2015), mais un ensemble complexe de communautés microbiennes interagissant avec leur hôte, le plus souvent eucaryote. Ces dernières années il a ainsi peu à peu été suggéré puis partiellement démontré que les communautés microbiennes des hôtes eucaryotes sont une composante fondamentale du fonctionnement de ces derniers (Schlaeppli & Bulgarelli, 2015, Turner et al., 2013, Berg et al., 2014a, Berg et al., 2014b). Les concepts de microbiote, microbiome et holobionte sont alors apparus et ont permis de renouveler le cadre conceptuel des approches d'écologie microbienne, mais aussi plus récemment des travaux de pathologies humaines, animales et végétales.

Le microbiote a été défini comme l'ensemble des microorganismes qu'ils soient eucaryotes, procaryotes ou viraux vivants dans un environnement spécifique. Cet environnement où vivent ces communautés de microorganismes est appelé microbiome. Les progrès technologiques ayant contribué à une meilleure caractérisation des communautés microbiennes et de leurs fonctions ont entraîné une constante réévaluation à la hausse du rôle du microbiote dans le fonctionnement des organismes hôtes (Müller et al., 2016, Villanueva-Millán et al., 2015, Engel & Moran, 2013, Ottman et al., 2012). Ces résultats assez récents ont questionné de nombreux travaux plus anciens obtenus par exemple en génétique fonctionnelle qui n'avaient jusqu'alors étudié que les gènes de l'hôte eucaryote pour définir ses capacités fonctionnelles. Les études successives du microbiote humain, ont montré qu'il existe environ 20 000 gènes humains et 2 à 20 millions de gènes microbiens (Wischmeyer et al., 2016) (Figure 1.7). Elles ont soulevé la question clé du rôle joué par le microbiote dans la santé et le développement de l'hôte et dans sa gestion potentielle pour améliorer les capacités fonctionnelles de l'hôte (Gopal & Gupta, 2016).

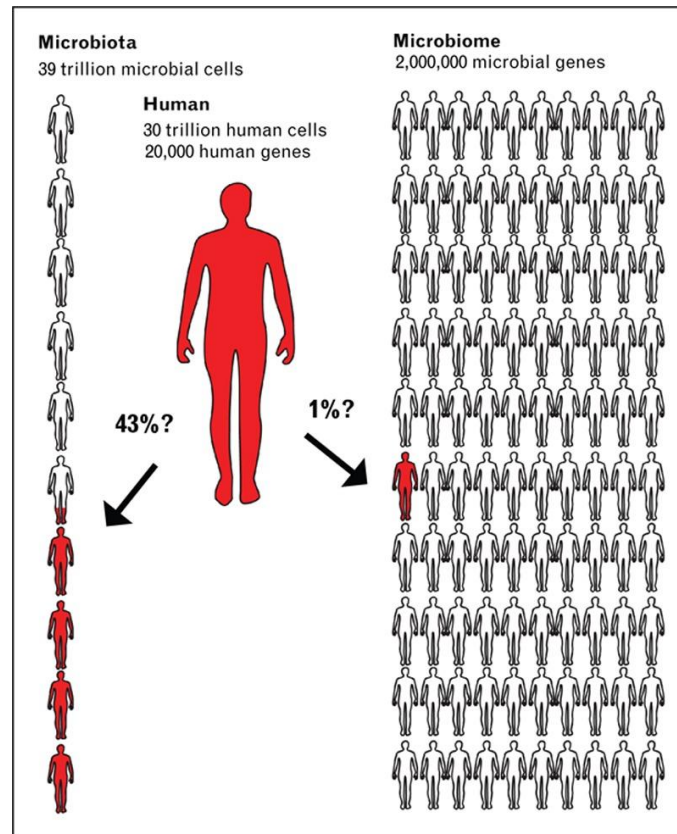


Figure 1.7 : Rapport entre les cellules et les gènes microbiens et humains dans le corps humain. (Wischmeyer et al., 2016)

Les plantes ne sont pas différentes des autres organismes et on estime que leurs microbiotes des plantes jouent un rôle important dans leurs croissances et leurs développements que ce soit par l'accès aux nutriments ou encore dans la résistance à des stress biotiques et abiotiques (Compant et al., 2019, Zahn & Amend, 2017, Busby et al., 2016, Castrillo et al., 2017). Des travaux ont spécifiquement démontré un très large éventail d'effets bénéfiques du microbiote sur leur hôte végétal comme la suppression de maladie (Mendes et al., 2011, Ritpitakphong et al., 2016), l'amorçage du système immunitaire de la plante hôte (Van der Ent et al., 2009), l'induction de la résistance systémique (Zamioudis et al., 2015), une augmentation de la disponibilité des nutriments (Van der Heijden et al., 2015), une augmentation de la tolérance à des stress abiotiques (Rolli et al., 2015), une adaptation à des variations environnementales (Haney et al., 2015), ou l'assistance à l'établissement d'associations mycorrhiziennes (Garbaye, 1994). Ce constat a amené à revoir certains éléments du fonctionnement et de l'évolution des organismes, en considérant les organismes et leurs microbiotes associés comme des supraorganismes. En 2002, le biologiste américain Forest Rohwer redéfinit le terme d'holobionte (du grec *holo*, tout, et *bios*, vie) créé par Lynn Margulis en 1991 (Margulis & Fester, 1991) comme étant un ensemble constitué de l'organisme hôte et de tous les microorganismes avec lequel il interagit

(Rohwer et al., 2002). De façon très intéressante, un article récent fait même remonter aux années 1940 et aux travaux du biologiste allemand Meyer-Abich (Baedke et al., 2020) le concept d'holobionte (« holobiosis ») et d'une théorie évolutive intégrant hôte et communautés microbiennes associées. Plus récemment, Rosenberg a proposé en 2011 le concept d'hologénome en complément du terme holobionte en se basant sur la synthèse de résultats de différentes études sur le microbiote décrivant les communautés microbiennes des écosystèmes marins, de l'homme et des plantes, (Rosenberg & Zilber-Rosenberg, 2011). Dans le concept d'hologénome, Rosenberg considère tous les génomes d'un holobionte comme une unité, conférant ainsi à l'organisme hôte des propriétés fonctionnelles représentatives de son holobionte. L'utilisation des concepts de l'holobionte et de l'hologénome remet ainsi en question la définition ontologique de l'individu. Un individu pourrait ainsi n'être pas uniquement défini par l'ensemble des cellules du « soi », mais également par l'ensemble des cellules du microbiote ou « non soi ». Les concepts d'holobionte et d'hologénome ont été extrêmement débattus ces dernières années. Partisans et sceptiques se sont opposés sur la pertinence de ce cadre éco-évolutif pour répondre à des questions sur la sélection ou la transmission du microbiote (Theis et al., 2016). Ce débat n'est pas limité qu'aux modèles humains et animaux. Le concept d'holobionte pour les plantes a aussi été largement débattu ces dernières années (Vandenkoornhuyse et al., 2015) et a stimulé la prise en compte de nouvelles questions (Vandenkoornhuyse et al., 2015) :

- Comment définir précisément l'holobionte végétal quand la composition du microbiote peut être différente durant le cycle de vie de l'hôte ? Comment mettre en évidence l'hétérogénéité du microbiote végétal au sein d'un organe donné lors des processus d'adaptation aux changements environnementaux ? Un microbiote cœur existe-t-il pour un hôte végétal donné de façon spatio-temporelle ?
- Comment distinguer les associations hôte/microbiote transitoires et durables ? Comment définir la part héritée de façon verticale à travers les graines ou acquise de façon horizontale post germination ?
- Quel est le compromis biologique qui permet à la fois à la plante de se défendre contre le microbiote endosphérique et a profité de ce dernier pour « augmenter » ses capacités immunitaires ?
- Enfin d'un point de vue évolutif et ontologique, comment envisager que l'unité de sélection soit l'holobionte ? Ce dernier point est particulièrement débattu et a fait l'objet de nombreuses opinions scientifiques contradictoires ces dernières années (Guerrero et

al., 2013, Osmanovic et al., 2018, Roughgarden et al., 2018, Gilbert, 2019, Madhusoodanan, 2019).

Nous focaliserons dans cette synthèse bibliographique sur l'état de l'art concernant « l'holobionte riz » constitué du riz comme hôte, et de son microbiote associé.

1.8. Le riz, une plante modèle

Parmi les 23 espèces que comprend le genre *Oryza*, *O. sativa* d'origine asiatique est l'espèce de riz la plus cultivée dans le monde devant l'espèce d'origine africaine *O. glaberrima*. (Ohyanagi et al., 2016). La diversité génétique et la structure des populations de l'espèce *O. sativa* a fait l'objet d'un très grand nombre d'études, s'appuyant sur des marqueurs biochimiques, moléculaires et génomiques. Le séquençage récent de 3010 génomes de l'espèce *O. sativa* a révélé sa très forte diversité génétique (Figure 1.8), notamment liée à l'origine géographique des variétés et corrélée aux capacités d'adaptation de la céréale (Wang et al., 2018).

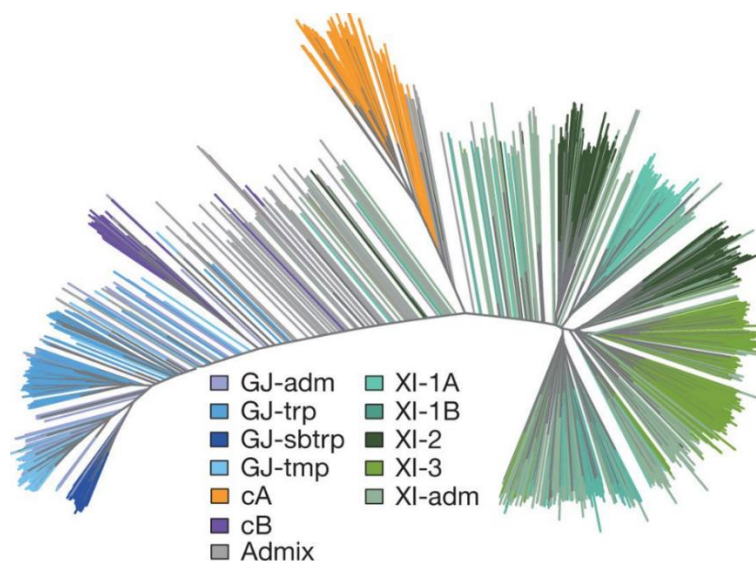


Figure 1.8 : Arbre représentant la diversité génétique de 3 010 échantillons de *Oryza sativa*. La reconstruction de l'arbre phylogénétique a été réalisée avec la méthode Unweighted neighbour-joining. Les échantillons sont colorés en fonction de leur affectation d'ADMIXTURE à $k = 9$. Les clusters XI correspondent à la diversité asiatique des riz *Oryza sativa indica* (XI-1A d'Asie de l'Est, XI-1B de variétés modernes d'origines diverses, XI-2 d'Asie du Sud et XI-3 d'Asie du Sud-Est). Les clusters GJ correspondent à la diversité asiatique des riz *Oryza sativa japonica* (GJ-tmp pour les japonica tempérés d'Asie de l'Est, GJ-sbtrp pour les japonica subtropicaux d'Asie du Sud-Est et GJ-trp pour les japonica tropicaux d'Asie du Sud-Est). Les groupes cA et cB correspondent à des génotypes de riz principalement retrouvés en Asie du Sud. Les termes adm et admix indiquent la présence d'amixture entre les génotypes indica et japonica (selon (Wang et al., 2018)).

Ces travaux ont, en outre, permis de distinguer des patrons complexes d'introggression observés dans les gènes domestiqués suggérant de multiples événements de domestication de l'espèce (Wang et al., 2018). La culture intensive du riz depuis la Révolution Verte a cependant entraîné une diminution de la diversité génétique des variétés cultivées dans les agrosystèmes rizicoles a été accompagnée d'une augmentation de l'impact des maladies et d'une augmentation de l'utilisation des intrants chimiques (Nalley et al., 2016). Alors que la consommation de riz devrait doubler d'ici 2050 (Wiebe, 2009), la gestion des maladies du riz et notre capacité à faire baisser leur impact sur les rendements est un enjeu crucial des travaux de recherche actuellement développés à l'échelle mondiale. En effet, les maladies les plus sévères (e.g. la pyriculariose, les bactérioses à *Xanthomonas* et la virose de la panachure jaune) peuvent constituer des contraintes majeures de la production du riz. Concernant la maladie majeure du riz au niveau mondial - la pyriculariose causée par le champignon *Pyricularia oryzae* - la sélection variétale qui permet l'utilisation de variétés résistantes reste aujourd'hui la principale alternative aux fongicides chimiques. Cependant, des contournements de ces résistances par une ou des populations de l'agent causal apparaissent quelques années après la diffusion aux champs des variétés (Ashkani et al., 2015) nécessitant la production continue de nouvelles variétés résistantes. Une meilleure compréhension du fonctionnement de l'holobionte du riz pourrait mener à la mise en place de méthodes de lutte complémentaires (voire alternatives) aux méthodes de lutte actuellement utilisées (luttés prophylactiques, curatives et génétiques) pour gérer les maladies du riz. Dans ce cadre-là, de nombreuses recherches ont été récemment initiées avec pour but de comprendre les dynamiques évolutives et écologiques du microbiote du riz et mieux identifier les facteurs biotiques et abiotiques impliqués dans la structuration du microbiote du riz. Ces travaux pourraient se traduire à court ou à moyen terme par la mise en place de méthodes de lutte basées sur le biocontrôle des microorganismes pathogènes du riz.

1.9. Les communautés microbiennes associées au riz

1.9.1. Les microbes colonisent différents compartiments de la plante

Tous les tissus de la plante hébergent différentes communautés microbiennes (Figure 1.9). On peut distinguer par exemple la phyllosphère (à la surface des tissus aériens), la rhizosphère (la zone étroite entourant et influencée par la racine), l'endosphère (qui représente l'intérieur des tissus), la spermosphère (Bulgarelli et al., 2013, Berg et al., 2014b, Turner et al., 2013).

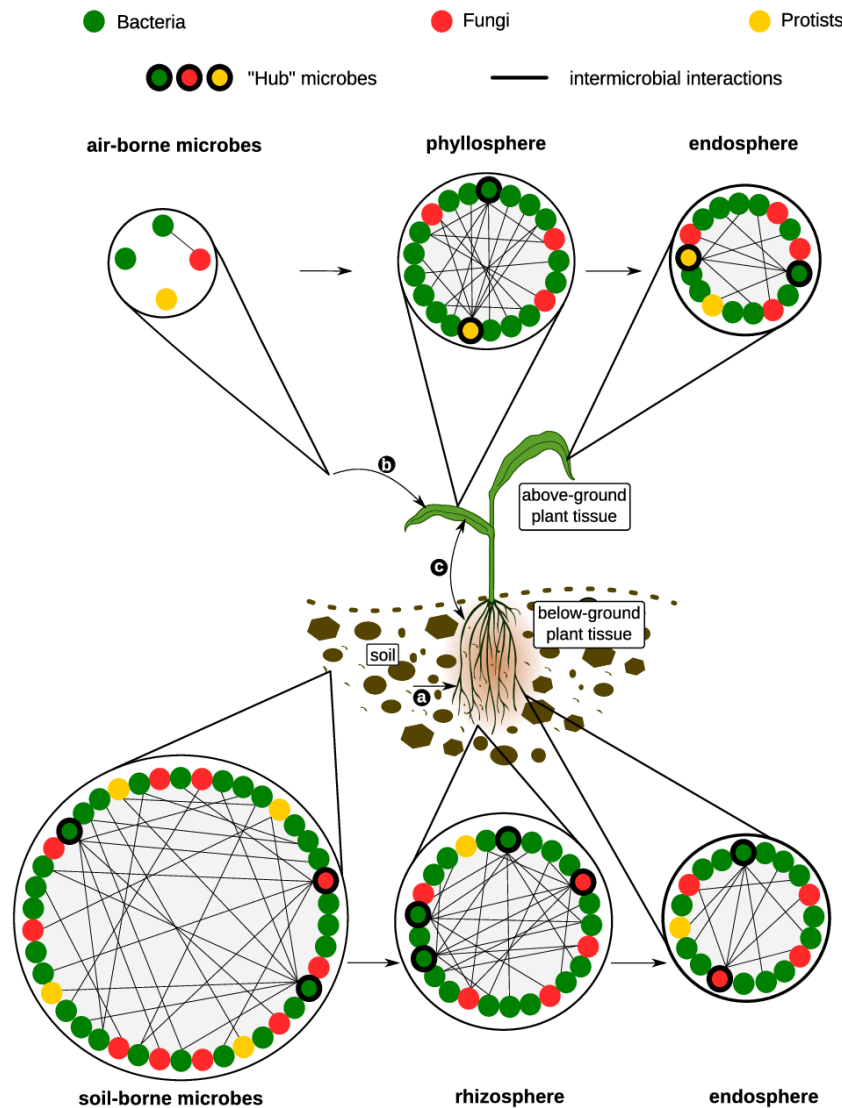


Figure 1.9 : Structure des communautés microbiennes dans différents compartiments végétaux. La figure illustre les communautés microbiennes dans le sol, l'air, la rhizosphère, la phyllosphère et à l'intérieur des tissus végétaux (endosphère). Cette figure illustre également les réseaux d'interactions entre les microorganismes. Dans chacun de ces habitats, les microorganismes (représentés par des cercles colorés) peuvent interagir positivement, négativement ou ne pas interagir pas avec d'autres microorganismes. Des microbes spécifiques, définis comme des espèces « clés » (cercles noirs en gras), sont fortement liés à d'autres microbes au sein des réseaux et exercent probablement une influence plus forte sur la structure des communautés microbiennes. (a) Les microbes associés aux racines proviennent principalement du sol. (b) Les microbes associés aux feuilles proviennent de diverses sources telles que les aérosols, les insectes ou la poussière. (c) Un déplacement des microorganismes du microbiote des racines et des feuilles peut être réalisé dans le compartiment endophyte (Hassani et al., 2018).

Ces environnements peuvent être influencés par différents facteurs biologiques et environnementaux. La rhizosphère est la zone qui est influencée par les exsudats racinaires. Lors de la rhizodéposition, les cellules racinaires sécrètent un grand nombre de composés (sucres, acides, vitamines, phytosidérophores, polysaccharides) qui vont attirer de nombreux microbes (Bulgarelli et al., 2013, Turner et al., 2013) . Même si certains composés nutritifs

comme des sucres peuvent être sécrétés au niveau des feuilles, la phyllosphère est considérée comme un milieu relativement pauvre, au niveau nutritif, comparée à la rhizosphère (Turner et al., 2013, Vorholt, 2012). La phyllosphère est par ailleurs un environnement plus dynamique que la rhizosphère du fait des fluctuations des conditions abiotiques moins tamponnées que dans la rhizosphère. Les microbes résidents de la phyllosphère doivent être capables de s'adapter aux variations de température, d'humidité et de radiations (Turner et al., 2013, Vorholt, 2012, Schlaeppi & Bulgarelli, 2015). La surface foliaire sur laquelle vont se déposer les microbes, tels que la cuticule foliaire qui définit le côté hydrophobe de la surface des feuilles, change avec l'âge. Les communautés microbiennes changent donc au cours des saisons (Lindow & Brandl, 2003). À la différence des microorganismes de la rhizosphère qui peuvent utiliser pour leurs croissances d'autres sources de carbone que celles fournies par la plante, les microorganismes de la phyllosphère vivent dans un milieu oligotrophe et sont pour la plupart très dépendants de leurs capacités à utiliser les quelques ressources en nutriments fournies par la plante.

L'endosphère comprend les tissus internes de la plante qu'ils soient aériens ou souterrains. La colonisation de ces tissus par les microbes est un processus actif (Reinhold-Hurek et al., 2015) où des gènes de la plante pourraient être impliqués dans la reconnaissance de bactéries bénéfiques (Reinhold-Hurek et al., 2015, Hardoim et al., 2015, Reinhold-Hurek & Hurek, 2011). Les microbes qui ont évolué pour coloniser ce milieu contraint ont acquis des propriétés spécifiques (Reinhold-Hurek et al., 2015, Hardoim et al., 2015). Il est vraisemblable que les endophytes et les rhizobactéries de la rhizosphère, jouent un rôle majeur comme bactéries promotrices de la croissance des plantes (PGPR pour Plant-Growth-Promoting-Bacteria) et comme agents de biocontrôle (Hardoim et al., 2015).

1.9.2. Les communautés microbiennes des compartiments racinaires

A ce jour, la composition des communautés microbiennes du compartiment racinaire est la plus étudiée (Kim & Lee, 2019). Les racines sont en relation avec le sol qui par sa structure physique supporte la plante et par sa composition chimique apporte à la plante les éléments nutritifs dont elle a besoin. Par ailleurs, le sol représente un des plus riches écosystèmes microbiens et contribue significativement à la biodiversité hors du sol et au fonctionnement des écosystèmes terrestres (Bardgett & van der Putten, 2014). Par exemple, les microbes du sol sont déterminants dans les cycles du carbone et de l'azote.

La structure d'une coupe de racines de riz est composée de trois compartiments principaux : la rhizosphère, le rhizoplan et l'endosphère (Figure 1.10). Les microbiotes de ces compartiments sont recrutés principalement à partir du sol environnant. La diversité microbienne décline séquentiellement du sol nu à la rhizosphère puis au rhizoplan et à l'endosphère de la racine suggérant une compétition croissante avec des habitats plus étroitement définis (Müller et al., 2016). Nous allons détailler les spécificités de chacun de ces compartiments et les raisons pour lesquelles les communautés microbiennes y ont des structures différentes.

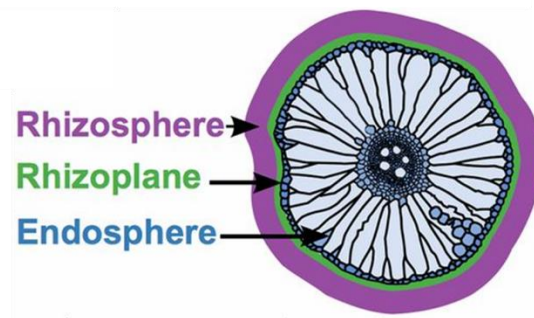


Figure 1.10 : Représentation d'une coupe transversale de racine de riz illustrant les différents compartiments hébergeant les communautés microbiennes. La rhizosphère est présentée en violet elle correspond au sol adhérent aux racines, le Rhizoplan en vert correspond à la surface des racines, l'endosphère en bleu correspond à la partie interne de la racine, (Edwards et al., 2015).

1.9.2.1. Le sol

Le sol des rizières est un milieu particulier. Il présente régulièrement des conditions d'anaérobiose en raison de l'irrigation ou de l'inondation permanente des parcelles de riz. Ces conditions anoxiques conduisent à des émissions de méthane conséquentes (Liesack et al., 2000). Cependant le transport de l'oxygène depuis les parties aériennes de la plante jusqu'à ses racines permet de créer des conditions aérobies au niveau de la rhizosphère (Figure 1.11). L'hétérogénéité de la concentration en oxygène dans la sphère racinaire influence la composition et la distribution spatiale des communautés bactériennes qui se répartissent suivant des microzones oxiques, oxiques-anoxiques ou anoxiques (Lee et al., 2015).

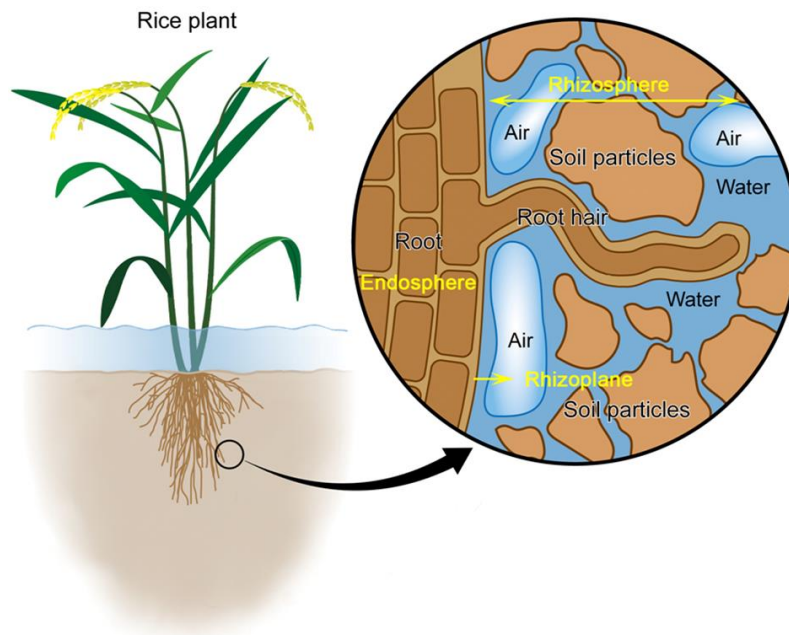


Figure 1.11 : Schéma d'une section de racine de riz présentant la structure de la rhizosphère, du rhizoplan et de l'endosphère dans un sol de rizière inondée. La sécrétion d'oxygène par les racines de riz entraîne la formation de zones oxygènes au niveau de la rhizosphère et d'un environnement anoxique au niveau du rhizoplan au niveau (Ding et al., 2019).

Ainsi, le sol des rizières peut donc être divisé en trois compartiments : la surface du sol oxygène, le sol nu anoxique et la rhizosphère et le rhizoplan oxygènes-anoxiques (Liesack et al., 2000). Une forte hétérogénéité dans le sol des rizières existe avec en particulier une hétérogénéité des conditions chimiques avec de fortes variations dans l'espace (Kim & Lee, 2019).

1.9.2.2. La rhizosphère

La rhizosphère bénéficie à la fois des apports en oxygène proposé par la plante et des exsudats de composés organiques qu'elle secrète (Achouak et al., 2019, Olanrewaju et al., 2019) (Figure 1.11). La libération par la plante de sources énergétiques au niveau de la rhizosphère du riz sous la forme de composés organiques, attire des microorganismes spécifiques au niveau de ce compartiment (Bacilio et al., 2003, Breidenbach et al., 2016). Il a été montré que les variations phénologiques de la plante au cours de son développement influence la composition de ses exsudats racinaires et entraîne des changements des communautés microbiennes de la rhizosphère (Aulakh et al., 2001, Lu et al., 2002). Le taux d'exsudation augmente en général jusqu'au stade floraison et décroît par la suite. L'attraction de certains microorganismes par la plante lui permet souvent d'en tirer à son tour un bénéfice en termes de croissance, de développement, de nutrition ou encore de résistance à des stress biotiques et abiotiques (Pili et

al., 2016, Andreo-Jimenez et al., 2019, Etesami & Alikhani, 2016, Lakshmanan et al., 2016, Kim & Lee, 2019).

La diversité et la composition des communautés bactériennes et fongiques racinaires varient avec le stade de développement du riz. Des différences de ces communautés ont été mis en évidence par électrophorèse sur gel en gradient dénaturant au stade tallage, au stade remplissage des grains et au stade maturité des grains (Hussain et al., 2011). Edwards et coll. ont montré un changement important de composition des communautés bactériennes du riz lors de la transition de la période végétative à la période reproductive, qui intervient dans leurs travaux autour de 7 à 8 semaines après germination (Edwards et al., 2018). Ces effets peuvent être corrélés aux variations d'exsudations racinaires (Aulakh et al., 2001).

Comme mentionné ci-dessus, la rhizosphère par ses caractéristiques renferme des microbes anaérobies stricts ou facultatifs et des microbes aérobies stricts. Elle est largement occupée par les protéobactéries (principalement les classes alpha, bêta et deltaprotéobactéries) ; les acidobactéries ; les actinobactéries et les chloroflexi. Les communautés bactériennes du riz majoritairement composées à partir de ces groupes bactériens sont relativement spécifiques et généralement distinctes de celles observées dans les rhizosphères d'autres plantes cultivées (Ding et al., 2019). Il est à noter en particulier dans la rhizosphère du riz, que la classe Deltaproteobacteria est nettement enrichie (Ding et al., 2019) et que les archées sont abondantes, incluant des archées méthanogènes et des archées non-méthanogènes (*Halobacteria*) présentes dans les couches superficielles du sol (Ma et al., 2019, Lee et al., 2015). Les communautés d'archées de la rhizosphère du riz sont communément composées de membres appartenant aux embranchements Crenarchaeota, Thaumarchaeota et Euryarchaeota. Ce type de composition en archées est spécifique de la rhizosphère du riz en comparaison des compositions d'archées retrouvées au sein des rhizosphères d'autres plantes cultivées où les membres de l'embranchement Thaumarchaeota sont généralement majoritaires (Ding et al., 2019). Chez les champignons, les communautés de la rhizosphère du riz se distinguent aussi de celle des rhizosphères d'autres plantes cultivées, notamment par la présence assez commune de membres appartenant à l'embranchement Chytridiomycota (Ding et al., 2019). Parmi les microorganismes de la rhizosphère du riz, il a été démontré que certains ont un effet bénéfique pour la plante. A titre d'exemple, des membres du genre *Aspergillus* (embranchement Ascomycota) ont la capacité d'induire la solubilisation du phosphore, alors que des champignons mycorhiziens arbusculaires (e.g. *Rhizophagus* spp.) peuvent faciliter le transport et l'assimilation du phosphore par les plantes (Mwajita et al., 2013, Mendes et al., 2013).

Cependant, les microorganismes colonisant la rhizosphère ne sont pas tous bénéfiques pour la plante. La caractérisation des communautés microbiennes permet, entre autres, de mettre en évidence des microorganismes pathogènes coexistant avec des microorganismes commensaux, et des microorganismes mutualistes pour la plante. Les maladies du riz causées par des microorganismes pathogènes sont moins présentes sur les parties racinaires que sur la partie aérienne. Toutefois, le sol des rizières peut servir de réservoir aux pathogènes qui peuvent infecter les plantes via leurs systèmes racinaires ; c'est le cas par exemple des champignons pathogènes du riz *Villosiclava virens* et *Ustilaginoidea virens* (Yong et al., 2018, Prakobsu & Ashizawa, 2017).

1.9.2.3. Le rhizoplan

Le rhizoplan représente la surface de la racine qui est colonisée par les microorganismes et forme la barrière entre rhizosphère et endosphère (Ding et al., 2019, Edwards et al., 2015). Seule une portion du microbiote de la rhizosphère se retrouve sur le rhizoplan avec des taxa enrichis par rapport à la rhizosphère (Edwards et al., 2015). Le microbiote du rhizoplan est plus similaire à celui de l'endosphère suggérant que le rhizoplan contribue à la sélection du microbiote de l'endosphère. On y retrouve des taxa enrichis appartenant aux embranchements Protéobactéries (alpha-, beta- et delta-), chloroflexi et Bacteroidetes, mais aux embranchements Fibrobactères et Spirochaetes liés à la dégradation de la cellulose et quelques archées méthanogènes associées à cette surface (Edwards et al., 2015, Knief et al., 2012). Par ailleurs, on retrouve associées au niveau du rhizoplan des bactéries antagonistes d'agents pathogènes ou d'autres contribuant à la disponibilité en phosphore ou à la fixation d'azote (Ding et al., 2019).

1.9.2.4. L'endosphère racinaire

La partie interne de la racine sera présentée ci-dessous dans le paragraphe dédié à l'endosphère du riz.

1.9.3. Les communautés microbiennes de la phyllosphère

La phyllosphère est définie comme étant la partie aérienne à la surface des plantes. Elle représente toutes les surfaces hors du sol comprenant généralement tiges et feuilles (Knief et al., 2012). Dans le cas du riz, on peut différencier les tiges de riz (la caulosphère), qui peut être en partie immergées, des feuilles. Les microorganismes appartenant à la phyllosphère sont épiphytes et majoritairement commensaux. Ils sont très difficiles à cultiver, ce qui a orienté les travaux de leurs caractérisations vers des approches de métagénomique (Kim & Lee, 2019).

Généralement, le compartiment de la phyllosphère est moins étudié que celui de la rhizosphère chez les plantes et cela est aussi vrai pour le riz.

Une étude analysant la phyllosphère du riz au stade de la floraison a reporté une dominance forte des embranchements des Alphaprotéobactéries (35%) et des Actinobactéries (38%). D'autres embranchements étaient représentés de façon moindre (< 5%), comme les Beta- et les Deltaproteobactéries, mais aussi quelques archées (Knief et al., 2012). Les auteurs de cette étude ont aussi montré que cette structure du microbiote de la phyllosphère était très différente de celle de la rhizosphère qui était décrite comme étant plus diverse et plus complexe (Knief et al., 2012). Une autre étude du microbiote de la phyllosphère basée sur des profils d'amplicons, a mis en avant l'importance des Gammaprotéobactéries. Une analyse récente des génomes de 3010 accessions de riz (Li et al., 2014) a permis d'extraire le métagénome bactérien et d'identifier certains facteurs associés à l'assemblage du microbiote de la phyllosphère (Roman-Reyna et al., 2019). Les Protéobactéries, Firmicutes et Actinobactéries dominent la phyllosphère. Sont retrouvées de façon plus marginales, quelques archées de l'embranchement Euryarchaeota qui inclut des méthanogènes (Roman-Reyna et al., 2019).

Comme pour les microorganismes de la rhizosphère du riz, il a été démontré que des microorganismes de la phyllosphère peuvent avoir un effet bénéfique pour le riz. Plusieurs études ont identifié des bactéries des genres *Methylobacterium* et *Rhizobium* à la surface de la phyllosphère, présentant des comportements mutualistes avec la plante hôte (Madhaiyan et al., 2009, Madhaiyan et al., 2004, Knief et al., 2012). Les bactéries du genre *Methylobacterium* apporteraient une protection contre les ultraviolets (UV) en constituant un biofilm protecteur (Vacher et al., 2016a). La capacité des bactéries du genre *Methylobacterium* à utiliser le méthanol comme source de carbone leur permet de bénéficier de la production d'alcools libérés par la plante à la surface de ses feuilles (Kurokawa et al., 2018). De plus, certains membres du genre *Methylobacterium* sont connus pour agir comme agents de bio-contrôle sur le riz contre *Rhizoctonia solani* (Madhaiyan et al., 2004). Des bactéries de ce même genre sont également connues pour promouvoir la croissance du riz (Madhaiyan et al., 2004, Chinnadurai et al., 2009). Les bactéries fixatrices d'azote du genre *Rhizobium* sont surtout décrites sur le riz pour promouvoir sa croissance, principalement par l'induction de la production de phytohormones (Nandi et al., 2009, Knief et al., 2012). Les communautés microbiennes de la phyllosphère du riz peuvent aussi présenter des microorganismes phytopathogènes. C'est par exemple le cas du champignon pathogène *Pyricularia Oryza* qui est l'agent causal de la Pyriculariose du riz, et du champignon pathogène *Rhizoctonia solani* qui entraîne des brûlures de la gaine, ou de

Xanthomonas oryzae causant deux maladies bactériennes du riz (Kim & Lee, 2019) . De nombreux travaux cherchant à isoler des agents antagonistes de ces microorganismes pathogènes par des approches de culture ont permis d'identifier des bactéries pouvant potentiellement servir d'agents de biocontrôle. Ces agents appartiennent aux genres bactériens *Bacillus*, *Streptomyces* ou *Erwinia* ou aux genres fongiques *Trichoderma*, *Alternaria* ou à différentes espèces de *Cladosporium* résidente de la phyllosphère du riz (Kim & Lee, 2019, Chaibub et al., 2020).

Bien que la rhizosphère et la phyllosphère soient des habitats très différents, certains organismes se retrouvent dans ces deux habitats. La présence de certains microorganismes dans ces deux habitats peut suggérer un déplacement de ces derniers dans la plante via le compartiment endophyte (Chi et al., 2005) ou par des contaminations à partir du sol (Bertani et al., 2016).

1.9.4. L'endosphère et les communautés microbiennes endophytes du riz

L'endosphère représente les parties internes de la plante, comprenant les tissus internes de la racine, la tige, les feuilles et les semences. Les communautés endophytes qui colonisent l'endosphère de la plante ont été initialement définies comme des microorganismes commensaux non-pathogènes. Ils sont originaires des communautés de microorganismes de la rhizosphère et de la phyllosphère ou proviennent des primo-endophytes présents dans la graine ou du matériel végétal de plantation (Hallmann et al., 2011). Même si les endophytes les plus communs sont des commensaux, on retrouve de façon moins fréquente des microorganismes mutualistes ou pathogènes de la plante hôte (Hardoim et al., 2015). Les communautés endophytes du riz arborent une diversité microbienne relativement importante, même si elle reste inférieure à celle de la rhizosphère et du rhizoplan (Ding et al., 2019, Edwards et al., 2015, Mano & Morisaki, 2008). Une étude de métagénomique des communautés endophytes des racines a mis en évidence une prédominance des membres de l'embranchement des Protéobactéries avec notamment des Gamma-Protéobactéries (incluant la famille des *Enterobacteriaceae*) et des Alpha-Protéobactéries (comprenant beaucoup de rhizobactéries abondantes) (Sessitsch et al., 2012). Parmi, les endophytes du riz, plusieurs sont considérées comme bénéfiques pour la plante. C'est le cas par exemple de *Pseudomonas stutzeri* A1501, initialement isolé de racines de riz en Chine (Vermeiren et al., 1999) qui a la capacité de fixer l'azote. D'autres PGPR endophytes associées au riz appartiennent au genre *Azorhizobium*. Ces bactéries sont connues pour produire des nodules artificiels, permettant la fixation de l'azote et la croissance des plantes (Tkacz & Poole, 2015). Par ailleurs, plusieurs bactéries endophytes

(dont certaines isolées des semences de riz) exhibent de fortes activités antifongiques contre des pathogènes du riz (Mano & Morisaki, 2008). Des champignons sont également présents dans les communautés microbiennes mutualistes du riz avec par exemple les champignons endomycorhiziens connus pour favoriser la croissance du riz par le transport du phosphate inorganique du sol dans les cellules de la plante (Paszkowski et al., 2002). Leur présence en symbiose avec les racines de riz améliore la résistance des plantes aux stress biotiques et abiotiques (Mbodj et al., 2018). Les champignons endomycorhiziens ont pour particularité d'être des champignons symbiotes obligatoires, c'est-à-dire incapables de survivre sans les racines de la plante hôte. Parmi ces champignons endophytes, certains montrent des capacités antagonistes vis à vis d'autres champignons pathogènes (Naik et al., 2009).

L'adaptation à l'endosphère et les capacités de ces microorganismes endophytes à vivre dans cette niche particulière sont peu connues. L'étude métagénomique du microbiote endophyte racinaire, a mis en évidence des protéines des systèmes de sécrétion nécessaires pour la translocation entre le cytoplasme et les membranes externes (Sessitsch et al., 2012). Il a été mis en évidence chez ces endophytes, des enzymes de dégradation des polymères de la plante, des gènes codant des protéines potentiellement impliquées dans la biosynthèse de sidérophores, une forte diversité de régulateurs transcriptionnels et des systèmes de communication par des molécules autoinductrices entre les bactéries (Sessitsch et al., 2012). Des gènes suggérant l'adaptation métabolique des endophytes du riz à leur micro-environnement avec un possible lien avec un gradient de concentration d'oxygène, ont aussi été révélés dans cette étude (Sessitsch et al., 2012). Enfin, des caractéristiques pouvant affecter la croissance et le rendement des plantes (production de phytohormones), des enzymes permettant la réalisation quasi complète du cycle de l'azote ont aussi été identifiés dans le métagénome endophyte des racines de riz (Sessitsch et al., 2012).

Par ailleurs, une étude de culturomique a permis d'isoler les communautés endophytes cultivables de l'endosphère racinaire, foliaire et de la tige à différents stades de développement du riz (Bertani et al., 2016). Parmi les 689 bactéries endophytes hypothétiques isolées des tissus internes, 53 ont montré de fortes propriétés bénéfiques de promotion de la croissance des plantes (fixation de l'azote, production d'auxines, activité anti-bactérienne ou anti-fongique) et 21 ont montré une capacité à coloniser l'endosphère via les racines (Bertani et al., 2016). Il a été montré que les microbes peuvent coloniser l'intérieur des racines de façon verticale par transmission des endophytes de la graine lors de la germination et par transmission horizontale où des microbes synthétisant des enzymes de dégradation des parois végétales peuvent

pénétrer, soit par les racines latérales émergentes ou accidentellement par des blessures (Ding et al., 2019, Mano & Morisaki, 2008).

1.9.5. Les communautés microbiennes de la graine

Des microorganismes peuvent être transmis par les graines. Ces dernières servent ainsi de vecteur privilégié à la transmission transgénérationnelle de microorganismes potentiellement bénéfiques (mais aussi potentiellement nuisibles) (Frank et al., 2017, Truyens et al., 2015). Les graines qui renferment l'embryon de la plante gouvernent la reproduction, la dissémination et aussi, lors de conditions défavorables, la survie des spermaphytes. La contribution de la transmission verticale du microbiote de la graine à la composition de la communauté microbienne de la future plante reste cependant globalement méconnue (Shade et al., 2017, Berg & Raaijmakers, 2018). Le microbiote peut coloniser la graine par les connexions vasculaires des différents tissus de la plante mère jusqu'à l'endosperme ou être transféré directement par les gamètes, par la fleur (glande à nectar, style) ou par le fruit (Nelson, 2018, Truyens et al., 2015, Compant et al., 2011). Ce microbiote pourrait avoir un rôle important pour protéger la plantule lors de la germination (Shade et al., 2017).

Du côté du riz, le rôle des graines comme source de bactéries endophytes reste cependant controversé (Mano & Morisaki, 2008), plusieurs travaux ont suggéré la transmission verticale via les graines, de souches bactériennes qui peuvent agir comme des agents de biocontrôle pour réduire l'effet d'agents pathogènes (Zhang et al., 2019, Cottyn et al., 2009, Cottyn et al., 2001, Bacilio-Jiménez et al., 2001, Mukhopadhyay et al., 1996) ou pour faciliter la croissance des jeunes plantules de riz (Verma et al., 2017). Le microbiote des semences semble plus conservé et moins variable que les microbiotes des autres compartiments de la plante. Différentes espèces bactériennes et fongiques ont été décrites à partir de semences de riz dont la flore microbienne de surface a été désinfectée (Kaga et al., 2009, Mukhopadhyay et al., 1996, Kim & Lee, 2019). Une présence systématique de quelques taxa appelés microbiote cœur a été reportée chez des semences de riz non désinfectées, produites dans différentes localisations géographiques ou conditions environnementales (Eyre et al., 2019). Ces taxa appartiennent principalement à la famille des *Enterobacteriaceae* avec les genres *Pantoea* et *Enterobacter* (Eyre et al., 2019). Des bactéries des genres *Pseudomonas*, *Acinetobacter* et *Xanthomonas* y sont également souvent retrouvés (Hardoim et al., 2012, Walitang et al., 2018, Zhang et al., 2019).

La transmission des agents pathogènes par les semences est un moyen important de dispersion des maladies (Gitaitis & Walcott, 2007). Les semences de riz peuvent ainsi jouer un rôle

important dans la transmission d'agents pathogènes et dans l'émergence de maladies dans de nouveaux agrosystèmes (Manandhar et al., 1998, Sakthivel et al., 2001).

1.10. Assemblage des communautés microbiennes du riz

1.10.1. Origine du microbiote du riz

Les microorganismes associés aux riz peuvent être acquis par différentes voies. Bien que minoritaires, certains microorganismes du microbiote de la plante sont transmis entre générations par voie verticale, via les semences par exemple (Clay & Schardl, 2002, Vannier et al., 2018, Frank et al., 2017). Ces microorganismes sont alors au contact de la plantule dès la germination et peuvent ou non se maintenir tout au long du cycle de développement de la plante (Shade et al., 2017, Wang et al., 2016). Même si la contribution de cette transmission verticale reste probablement relativement faible, nous avons vu dans le paragraphe précédent que ces quelques microbes issus de la graine pourraient avoir un rôle important dans la colonisation de la plantule et donc dans l'assemblage du microbiote du riz (Vannier et al., 2018).

L'acquisition des communautés microbiennes chez le riz, comme chez les autres plantes, se fait vraisemblablement majoritairement par voie horizontale (Frank et al., 2017). Le sol est ainsi le principal réservoir du microbiote racinaire (Mano & Morisaki, 2008, Breidenbach et al., 2016, Edwards et al., 2015) et des microorganismes symbiotiques du riz (Zhang et al., 2018). Cependant, d'autres sources d'inoculum peuvent participer à l'assemblage des communautés microbiennes, dont l'air et l'eau (Almaguer et al., 2012, Ciancio & Mukerji, 2008). Beaucoup de microorganismes pathogènes du riz ont été décrits pour avoir été transmis à partir de ces deux réservoirs environnementaux (Almaguer et al., 2012, Ciancio & Mukerji, 2008, Katsantonis et al., 2017). La colonisation des tissus aériens de la plante à partir de sources d'inoculum d'origines environnementales place souvent les microorganismes dans un nouvel environnement oligotrophe où les symbiontes ont de la difficulté à s'établir. Les microorganismes retrouvés à la surface des feuilles ne sont pas forcément adaptés à cet environnement et peuvent se retrouver sur la surface des feuilles de façon transitoire et passive sans être capables de s'y multiplier.

1.10.2. Dynamique de l'assemblage du microbiote du riz

L'âge de la plante ou son stade de développement semblent être des paramètres importants dans la structuration du microbiote des compartiments racinaires du riz (Hussain et al., 2012,

Edwards et al., 2015, Edwards et al., 2018, Imchen et al., 2019, Zhang et al., 2018). Lors d'une expérimentation, des plantules stériles de variétés de riz *japonica* ont été transplantées dans le sol d'un champ pendant des périodes allant de 0 à 13 jours (Edwards et al., 2015). Dès le premier jour après la transplantation, des bactéries ont été détectées par une approche de metabarcoding dans l'endosphère des racines ce qui montre une acquisition très rapide des bactéries endophytes par la plante. La proportion des bactéries a augmenté pendant les deux semaines de l'étude (Edwards et al., 2015). D'autres travaux ont analysé le microbiote sur tout un cycle de culture du riz, et il a été montré que la composition du microbiote racinaire variait avec le stade de développement des plantes (Edwards et al., 2015, Edwards et al., 2018, Imchen et al., 2019, Zhang et al., 2018) (Figure 1.12). Le microbiote racinaire a été très dynamique pendant la croissance végétative et on a observé un changement du microbiote du rhizoplan et de l'endosphère sept à huit semaines après la germination qui correspond à la transition vers la période reproductive de la plante. Le microbiote du rhizoplan et de l'endosphère se stabilise alors, mais pas celui de la rhizosphère (Edwards et al., 2018).

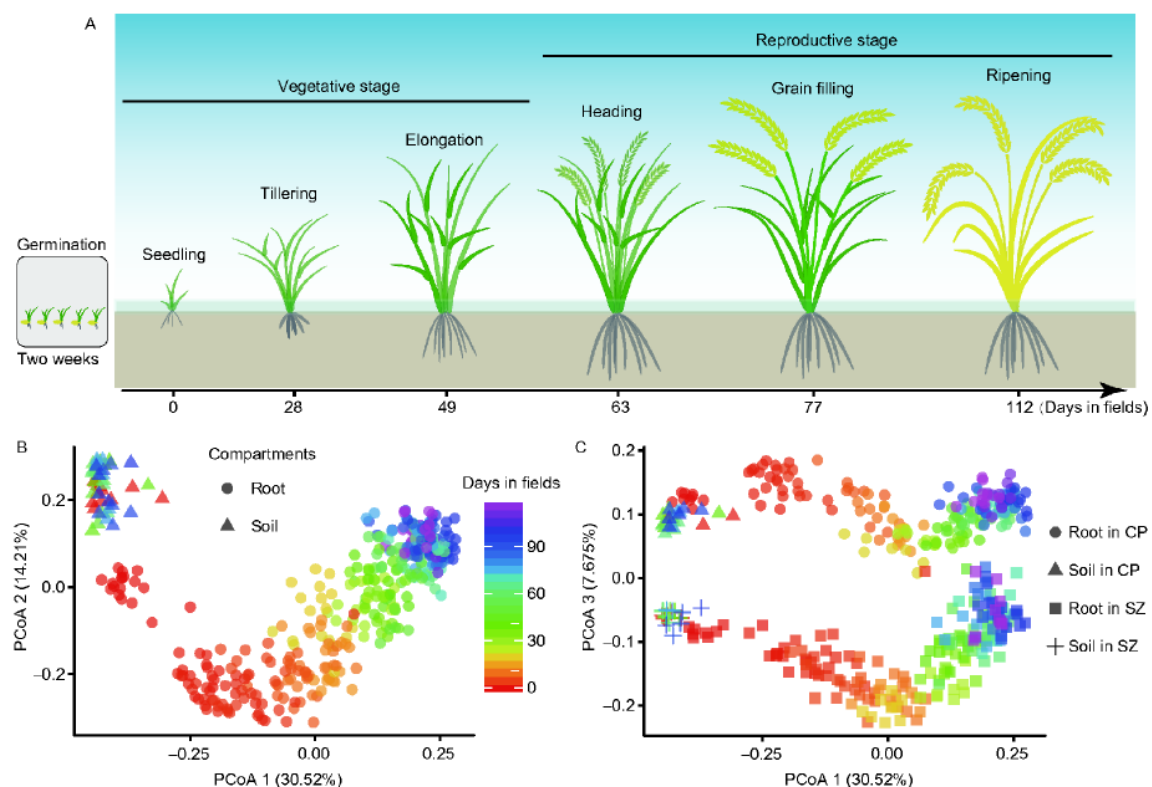


Figure 1.12 : Stade de développement du riz et dynamique de de l'assemblage du microbiote racinaire du riz. A) Diagramme du cycle de croissance du riz représentant les principaux stades de développement. B-C) représente l'analyse en coordonnées principales montrant que la structure du microbiote des racines du riz évolue au cours du temps et diffère du microbiote du sol. Les codes CP et SZ sur le graphique C) correspondent à deux champs de riz différents. (Zhang et al., 2018)

1.10.3. Les facteurs influençant la composition et la structure du microbiote

Bien que toutes les études précédentes s'accordent sur une dynamique de structuration des communautés microbiennes au cours du développement du riz, les différences de résultats entre les études ont pointé la complexité des assemblages microbiens et ont suggéré la présence d'autres facteurs de structuration biotiques et abiotiques (Edwards et al., 2015, Edwards et al., 2018, Imchen et al., 2019, Zhang et al., 2018) (Figure 1.13). Une connaissance de ces facteurs structurant les communautés microbiennes est nécessaire pour comprendre l'importance et le rôle des associations entre la plante hôte et ses microorganismes associés. Parmi ces facteurs, le compartiment et le génotype de la plante ainsi que l'environnement de celle-ci sont les plus documentés.

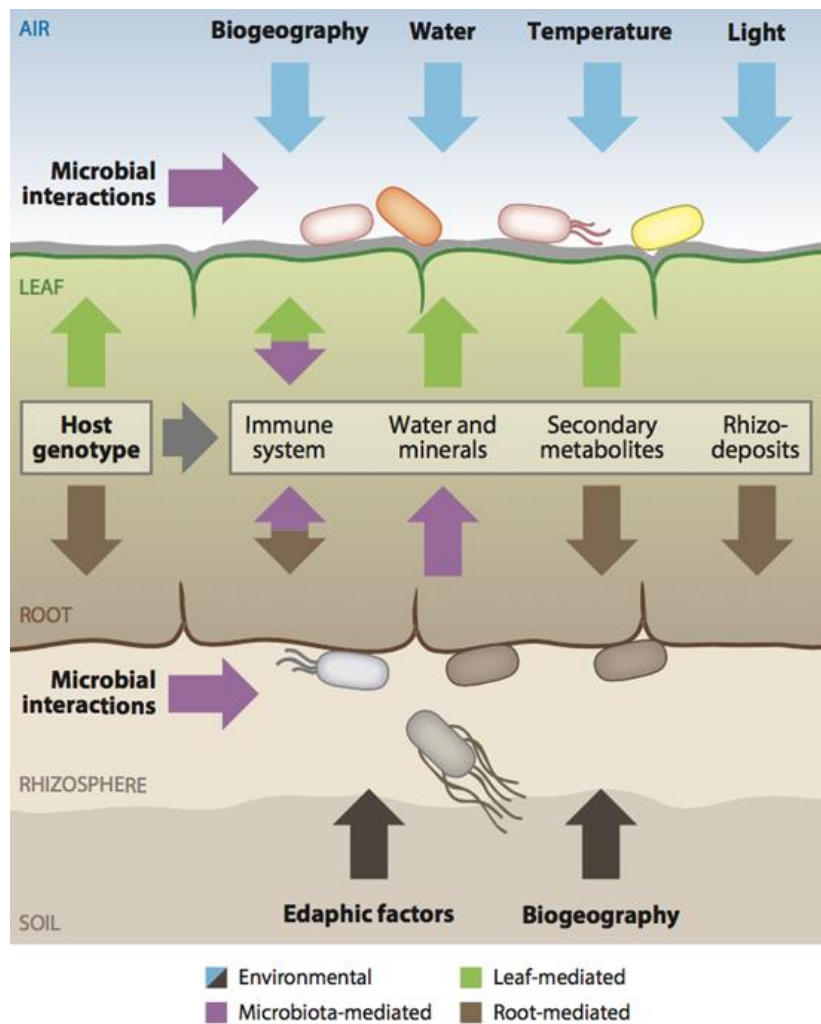


Figure 1.13 : Schéma présentant les facteurs biotiques et abiotiques influençant la structure des communautés microbiennes des plantes (Müller et al., 2016)

1.10.3.1. Le compartiment de la plante

Les compartiments végétaux que ce soit la rhizosphère, le rhizoplan, l'endosphère, ont tous un microbiote spécifique. Le compartiment de la plante apparaît comme un des facteurs influençant le plus la structure du microbiote (Yuan et al., 2011, Knief et al., 2012, Edwards et al., 2015, Edwards et al., 2018). Cette observation suggère une relation forte entre les microenvironnements créés par la plante autour de ses différents compartiments comme par exemple via l'exsudation de composés organiques ou des réactions immunitaires qui régulent la présence de certains microorganismes (Bacilio et al., 2003, Sasse et al., 2018). La richesse bactérienne apparaît comme plus importante dans la rhizosphère et le rhizoplan que dans les compartiments endophytes et de la phyllosphère du riz (Kim & Lee, 2019). L'extraordinaire biodiversité microbienne dans le sol et la forte concentration en composés organiques sécrétés par la plante dans le sol proche des racines du riz en fait une zone propice à la croissance bactérienne (Bertani et al., 2016). Les compartiments racinaires sont dominés par les communautés bactériennes alors que les compartiments aériens semblent plus dominés par les communautés fongiques (Chi et al., 2005, Wang et al., 2016).

1.10.3.2. Le génotype du riz

Les plantes peuvent former leur microbiote en recrutant certains microorganismes ou en filtrant certains endophytes en fonction de leur génotype. Le génotype de riz ne semble toutefois pas avoir d'effet fort sur la structure des communautés des compartiments racinaires lorsque ceux-ci sont génétiquement proches et font partie d'une même sous espèce d'*Oryza sativa* (Moronta-Barrios et al., 2018, Edwards et al., 2015, Edwards et al., 2018). *A contrario*, il a été observé des différences significatives de composition des communautés microbiennes entre les variétés de riz cultivés et les variétés de riz sauvages ou entre plusieurs sous-espèces de riz (Hardoim et al., 2011, Edwards et al., 2015, Edwards et al., 2018, Shenton et al., 2016, Andreo-Jimenez et al., 2019, Imchen et al., 2019, Shi et al., 2019). Certains groupes de microbes peuvent être spécifiquement affectés par le génotype. Il a été par exemple montré que plusieurs génotypes de riz ont des effets contrastés sur les populations de diazotrophes associées aux racines exprimant le gène *nifH* de la nitrogénase (Knauth et al., 2005). Tous ces travaux ont étudié exclusivement l'effet du génotype sur le microbiote racinaire du riz. Un travail récent a cependant montré que les communautés microbiennes de la phyllosphère du riz entre plusieurs sous-espèces d'*Oryza sativa* étaient également en partie structurées par le génotype de la plante (Roman-Reyna et al., 2019). Une autre analyse par polymorphisme de longueur de fragments

de restriction terminaux (RFLP-T) a révélé un effet du génotype (variétés de type *indica*) sur la diversité du microbiote endophyte des semences de riz (Walitang et al., 2018).

1.10.3.3. Les pratiques culturales

Les diversités α et β du compartiment racinaire du riz varient significativement avec les pratiques culturales, révélant une plus grande diversité de la rhizosphère dans une culture conventionnelle avec intrants chimiques que dans une culture biologique, mais aucune différence n'a été observée au niveau du microbiote du rhizoplan ou de l'endosphère (Edwards et al., 2015). La diversité du microbiote de la phyllosphère est aussi modifiée par différentes doses de fertilisation azotée (Thapa et al., 2018).

1.10.3.4. Les autres facteurs environnementaux

D'autres paramètres environnementaux participent à la structuration des communautés microbiennes du riz. A partir de l'analyse de génomes de riz, il a par exemple été montré que la diversité du microbiote de la phyllosphère dépend de l'origine géographique des plantes (Roman-Reyna et al., 2019). Des facteurs abiotiques peuvent aussi modifier le microbiote. En période de manque d'eau, le microbiote de l'endosphère du riz est plus affecté que les autres compartiments racinaires (Andreo-Jimenez et al., 2019, Santos-Medellín et al., 2017, Edwards et al., 2018). De même des semences de riz issues de variétés exprimant différents niveaux de tolérance à la salinité ont des microbiotes endophytes qui répondent différemment quand leurs plantes hôtes sont exposées à différentes concentrations de sel dans le sol (Walitang et al., 2018). Les changements de communautés microbiennes observés peuvent être une conséquence de ces facteurs environnementaux, mais cela peut aussi permettre d'expliquer les variations de sensibilités des plantes à ces facteurs.

Beaucoup d'études aux champs se placent dans des conditions homogènes à un temps donné, mais ne prennent pas en compte des facteurs indirects qui pourraient avoir un effet sur la structure des communautés microbiennes associées à la plante. L'historique d'utilisation des sols peut avoir des effets sur les communautés microbiennes des sols (Silva Parra et al., 2018) et indirectement structurer les communautés microbiennes associées au riz pour les années suivantes (Edwards et al., 2019, Edwards et al., 2015)

1.11. Etude du virome du riz

Les virus du riz ou une partie de ces virus n'ont jamais à notre connaissance été pris en compte dans les études du microbiote du riz réalisées ces dernières années. Par ailleurs, aucune étude n'a caractérisé à ce jour le virome complet du riz par des approches de métagénomique. De nombreux travaux de recherches ont permis de diagnostiquer et de caractériser de façon ciblée des virus pathogènes du riz, mais aucune étude n'a recherché de façon systématique et non ciblée tous les virus présents dans des échantillons de riz.

Une trentaine de virus pathogènes du riz ont été décrits à ce jour par des approches moléculaires classiques, ils appartiennent à plusieurs familles virales (Moriyama et al., 1999a, Uehara-Ichiki et al., 2013) :

- Famille *Phenuiviridae* avec rice stripe virus (RSV), rice hoja blanca virus (RHBV) et rice grassy stunt virus (RGSV)
- Famille *Reoviridae* avec rice dwarf virus (RDV), rice gall dwarf virus (RGDV), rice black streaked dwarf virus (RBSDV), rice ragged stunt virus (RRSV) et Southern rice black streaked dwarf virus (SRBSDV)
- Famille *Rhabdoviridae* avec rice transitory yellowing virus et rice yellow stunt virus (RYSV)
- Famille *Caulimoviridae* avec rice tungro bacilliform virus (RTBV)
- Famille *Secoviridae* avec rice tungro spherical virus (RTSV)
- Famille *Solemoviridae* avec rice yellow mottle virus (RYMV)
- Famille *Geminiviridae* avec maize streak virus (MSV), rice latent virus (RLV)
- Famille *Benyviridae* avec rice stripe necrosis virus (RSNV)
- Famille *Endornaviridae* avec oryza rufipogon alphaendornavirus (OrEV) et oryza sativa alphaendornavirus (OsEV)
- Famille *Potyviridae* avec rice necrosis mosaic virus (RNMV)

Par ailleurs le riz héberge un cortège de bactéries et de champignons qui peuvent être respectivement infectés par des phages et des mycovirus. Plusieurs travaux récents ont focalisé sur les mycovirus et sur leurs potentielles actions de contrôle de certains champignons pathogènes du riz (Moriyama et al., 2018, Urayama et al., 2010, Zheng et al., 2019). Il a ainsi été montré que des mycovirus appartenant aux genres *Endornavirus* et *Chrysovirus* induisent des phénomènes d'hypovirulence des champignons du riz qu'ils infectent (respectivement

Rhizoctonia solani (Zheng et al., 2019) et *Pyricularia oryzae* (Moriyama et al., 2018, Urayama et al., 2010). Des travaux similaires ont par ailleurs montré que des bactériophages peuvent avoir une action de contrôle de pathogènes bactériens du riz (Ranju et al., 2018, Buttimer et al., 2017). A titre d'exemple, des plantules issues de semences de riz ayant été traitées avec le phage ϕ XOF4 à une concentration de 10^8 PFU (Unité de Formation de Plaque) /ml se sont révélées être moins affectées en termes de symptômes causés par *Xanthomonas oryzae* que les plantules issues de semences non traitées par le même phage (Ranju et al., 2018).

1.12. Problématique et objectifs

Nous avons vu dans la revue bibliographique que les microorganismes en association avec un hôte peuvent étendre ses capacités fonctionnelles et ainsi contribuer à sa croissance et son développement ainsi qu'à son bon état de santé (Berg & Koskella, 2018, Vannier et al., 2019, Ritpitakphong et al., 2016, Castrillo et al., 2017). Il en résulte qu'aujourd'hui l'étude du fonctionnement d'un organisme végétal ou animal ne peut s'affranchir de l'étude de la diversité des microorganismes qui l'habitent et du rôle fonctionnel que ces derniers jouent au sein de cet holobionte constitué. Toutefois, les processus d'assemblage du microbiote des plantes et des forces évolutives qui gouvernent ces processus restent loin d'être compris et maîtrisés. A titre d'exemple, les rôles joués par les facteurs déterministes biotiques et/ou abiotiques et les processus stochastiques dans les dynamiques d'assemblage du microbiote ont été jusqu'à ce jour peu étudiés et quantifiés du fait de la difficulté d'identifier les nombreux effets confondants associés à ce type d'analyse.

L'objectif principal de ce travail de thèse a été, dans le cadre général d'une meilleure compréhension de l'assemblage du microbiote des plantes, de déterminer si la structure des communautés microbiennes associées à une plante peut être influencée par le génotype de celle-ci. Nous avons dans ce cadre général cherché à caractériser trois composantes majeures du microbiote des plantes, à savoir les virus, les bactéries et les champignons. L'analyse des virus des plantes n'a à notre connaissance jamais été réellement intégrée aux études du microbiote des plantes, majoritairement focalisées sur les bactéries et/ou les champignons.

Parallèlement aux travaux initiés sur le microbiote des plantes, des études ont récemment montré l'importance de la diversité génétique des hôtes cultivés (mélanges variétaux, cultures intercalaires, couverture des sols de vergers, etc.) pour le bon fonctionnement des agrosystèmes (Bardgett & van der Putten, 2014, Hooper et al., 2005). Ces itinéraires culturels basés sur la

diversité intra-spécifique et/ou inter-spécifique des plantes cultivées font actuellement figure d'alternatives aux systèmes agricoles intensifs issus de la Révolution Verte qui ont été orientés vers une simplification de la diversité des plantes cultivées, contribuant à une augmentation de la pression des microorganismes phytopathogènes. Deux voies sont possibles pour étudier l'impact de la diversité génétique des hôtes cultivés sur le bon fonctionnement des agrosystèmes : soit mettre en place des dispositifs expérimentaux *ad hoc* de terrain, soit explorer des agrosystèmes séculaires encore en place qui n'ont pas été affectés par les méthodologies employées lors de la Révolution verte.

Nous avons choisi d'aborder la question du rôle joué par le génotype de la plante cultivée en explorant le modèle riz, dont les diversités génétique et génomique étaient déjà bien caractérisées d'un point de vue inter-spécifique et intra-spécifique au début de cette étude (Wang et al., 2018). Sur ce modèle de plante, les travaux sur le microbiote du riz avaient alors ciblé essentiellement les procaryotes et mis en avant une dynamique rapide dans l'assemblage du microbiote, un effet fort du compartiment et des effets faibles du génotype et des pratiques culturales dans la structuration du microbiote du riz (Edwards et al., 2015, Knief et al., 2012). Des travaux s'étaient par ailleurs traduits par la caractérisation du microbiote endophyte associé au riz et par une meilleure compréhension du rôle fonctionnel de ce microbiote endophyte (Sessitsch et al., 2012). En revanche, peu de travaux avaient été réalisés sur la phyllosphère.

Nous avons cherché à comprendre comment le microbiote du riz (bactéries, champignons et virus) répond aux variations génétiques de l'hôte dans deux contextes expérimentaux différents : au niveau d'un dispositif paysan dans un agrosystème séculaire chinois et au niveau d'un dispositif expérimental *ad hoc* de terrain en Camargue. Le premier cas est l'agrosystème des terrasses du Yuanyang (aussi appelé Honghe Hani rice terraces system (HHRTS) en anglais) dans la province du Yunnan en Chine qui est un agrosystème centenaire et durable exhibant une très forte diversité génétique de riz traditionnels. Cette forte diversité dans ces terrasses rizicoles a été associée à des services écosystémiques, notamment la régulation des agents pathogènes. Cet agrosystème voit depuis plusieurs années une introduction notable de variétés de riz modernes. Nos hypothèses étaient que (i) cette introduction allait modifier la diversité génétique du riz dans cet agrosystème du HHRTS et (ii) que ce changement de diversité chez l'hôte allait impacter la structure du microbiote de l'hôte. Dans la deuxième situation qui était une parcelle expérimentale du Centre Français du Riz localisée en Camargue, nous avons analysé l'impact du génotype du riz sur son microbiote associé au cours du développement du riz.

Les questions spécifiques suivantes ont été posées dans ce travail de thèse :

- Quelle est la diversité virale, bactérienne et fongique du microbiote du riz dans les agroécosystèmes chinois et français ?
- Existe-t-il un microbiote cœur dans ces deux agrosystèmes rizicoles ?
- Quelle est la dynamique de structuration des communautés microbiennes (virus et bactéries endophytes) du riz dans l'agrosystème de Camargue ?
- Quel est l'effet du génotype du riz sur le microbiote du riz dans les deux agrosystèmes rizicoles ?
- Est-ce que la structure et la composition du microbiote sont modifiées par la présence de microorganismes pathogènes ?

Le manuscrit de thèse est composé de deux chapitres, le premier est axé sur les résultats obtenus à partir des échantillons récoltés en Chine et le deuxième focalise sur les travaux réalisés en Camargue. Ces deux chapitres seront suivis d'une discussion générale tentant de synthétiser les résultats obtenus dans les deux agrosystèmes.

Chapitre 2 : Caractérisation de l'holobionte du riz dans un agrosystème chinois centenaire et durable

2.1. Introduction générale du chapitre 2

Comme précédemment mentionné dans la problématique générale de la thèse, nous avons choisi dans ce premier chapitre, de travailler sur l'agrosystème en terrasses de la région chinoise du Yuanyang ou aussi appelé en anglais Honghe Hani rice terraces system (HHRTS). Cultivé depuis environ 1300 ans par la minorité ethnique Hani, cet agrosystème rizicole centenaire est célèbre pour ses paysages (Figure 2.1) et pour son exceptionnelle diversité de variétés traditionnelles de riz, (Zhu et al., 2000). Des inventaires récents ont dénombré dans cette région la présence d'au moins 195 landraces (terme anglais pour cultivars primitifs) de l'espèce domestiquée *Oryza sativa* et 47 landraces de riz sauvages des espèces *Oryza rufipogon* et *Oryza nivara* (Jiao et al., 2012).



Figure 2.1 : Localisation et présentation de l'agrosystème rizicole en terrasses de l'ethnie Hani de la vallée de Honghe.

Cet agrosystème est considéré comme durable. Les maladies rizicoles y ont peu d'impact et les rendements sont globalement bons (5 à 7 tonnes/ha). A titre d'exemple, des travaux ont montré une relation entre la forte diversité rizicole de cet agrosystème et la faible émergence d'épidémies de pyriculariose (Liao et al., 2016). Cependant, l'introduction récente de variétés modernes à haut rendement de riz pourrait perturber l'exceptionnelle durabilité de ce système.

Le remplacement de variétés traditionnelles de riz par quelques variétés modernes à haut rendement, risque en effet de globalement réduire la diversité génétique du riz et de modifier les itinéraires culturels ancestraux, avec un usage plus important d'engrais de synthèse et de pesticides. Ces modifications pourraient aussi avoir un impact sur la composition du microbiote associé au riz. Il est ainsi à craindre que l'ensemble de ces bouleversements puisse se traduire

par une perte de la durabilité de l'agrosystème multi-centenaire du Yuanyang, notamment à travers le développement d'épidémies. Cette hypothèse a été testée dans une enquête sociologique confrontant les observations de paysans cultivant des variétés modernes ou traditionnelles. Les résultats de cette enquête ont révélé un impact plus important des maladies dans les villages ayant remplacé les variétés de riz traditionnelles par des variétés de riz moderne (Dedeurwaerdere & Hannachi, 2019).

Ce chapitre de thèse a eu pour objectif principal de prolonger les travaux de l'enquête sociologique menée par Dedeurwaerdere et Hannachi (2019) en testant l'hypothèse que l'introduction de variétés modernes dans l'agrosystème du Yuanyang pourrait avoir un impact sur la composition du microbiote associé au riz. Nous avons choisi de focaliser nos travaux sur un village (Malizhai) de la région du Yuanyang connu pour faire coexister depuis quelques années les systèmes culturels traditionnels et modernes. Afin de nous assurer que nous pourrions tester l'hypothèse générale de ce chapitre, nous avons tout d'abord « géotypé » les plants de riz que nous avons récoltés. Les plants étaient identifiés comme appartenant à des variétés modernes ou traditionnelles par les paysans de la zone d'étude. Nous avons alors décidé d'analyser trois types de microorganismes du microbiote (virus, bactéries et champignons) des variétés de riz traditionnelles et modernes échantillonnées.

Ce chapitre de thèse est composé de trois sections : i) une analyse par métagénomique du virome des plants de riz modernes et traditionnels récoltés sur le site d'étude, ii) une étude approfondie de l'émergence dans l'agrosystème du Yuanyang du Southern rice black-streaked dwarf virus, un fijivirus détecté par nos travaux initiaux de métagénomique virale et iii) l'étude des communautés microbiennes (bactéries et champignons) associées aux plants de riz modernes et traditionnels récoltés sur le site d'étude.

2.2. Section 1 : Analyses du virome des variétés de riz traditionnelles et modernes cultivées au sein du village Malizhai

2.2.1. Matériel et méthodes

2.2.1.1. *Site d'étude et échantillonnage*

Une analyse de la diversité virale par métagénomique a été réalisée à partir d'un échantillonnage effectué en juillet 2016 provenant du village de Malizhai, situé dans la région chinoise du Yuanyang (Figure 2.2). Nous avons collecté des échantillons de riz dans 11 parcelles décrites

par les paysans comme étant cultivées avec des variétés traditionnelles et 8 parcelles cultivées avec des variétés modernes. Pour chaque parcelle échantillonnée, les feuilles, les tiges et les racines de 10 plants de riz ont été prélevées individuellement, puis conservées par déshydratation à l'aide de chlorure de calcium (CaCl_2) selon la technique de Bos (Bos, 1977). Le descriptif complet du site d'étude est présenté dans les matériels et méthodes des deux sections suivantes de ce chapitre de thèse.

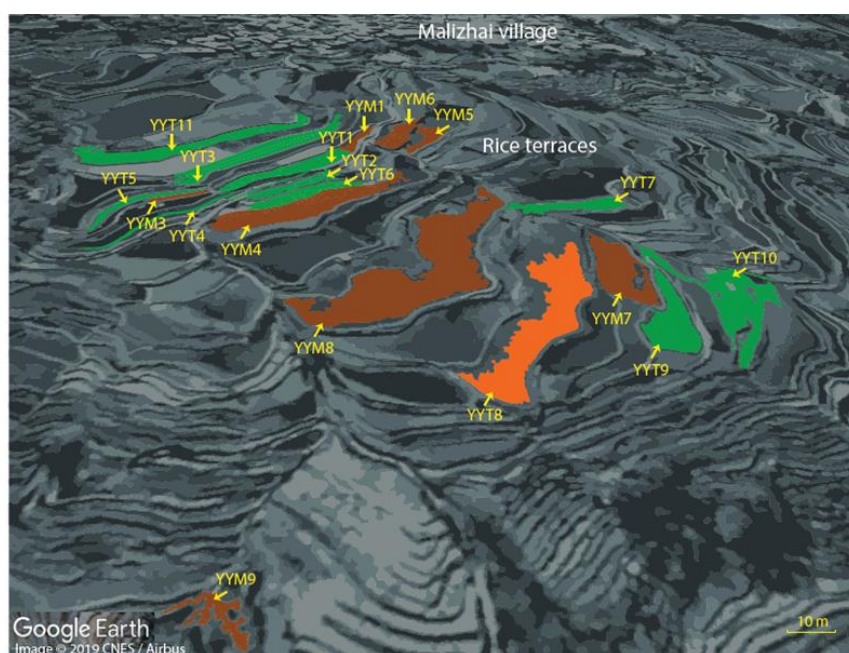


Figure 2.2 : Plan de la zone d'échantillonnage, les parcelles colorées correspondent aux parcelles échantillonnées ; les parcelles en marron sont celles cultivées avec des variétés modernes indica ; les parcelles en vert sont celles cultivées avec des variétés traditionnelles indica et la parcelle en orange correspond à une variété de riz traditionnelles japonica

2.2.1.2. *Semi-purifications virale et analyse du virome par métagénomique virale*

Les 190 feuilles récoltées dans les 19 parcelles ont été traitées individuellement. L'extraction et la semi-purification des particules virale ont été réalisées avec l'approche virion-associated nucleic acids (VANA), à partir de 0,3g de matière sèche pour chacun des échantillons, comme décrits dans le chapitre d'ouvrage technique rédigé par mon équipe d'accueil (François et al., 2018). Les produits d'amplification ont été séquencés par la technologie Illumina HiSeq 2500 qui permet le séquençage sens et anti-sens de fragments d'ADN sur une longueur totale de 150 paires de bases.

Les données de séquençage ont été démultiplexées avec l'algorithme Agrep (Ahmad et al., 2006) et nettoyées avec l'algorithme Cutadapt (Martin, 2011). Afin d'éliminer le plus possible de séquences nucléotidiques associées à la plante hôte et réduire le temps de calcul des étapes

suivantes, un mapping « end to end » des séquences (reads) a été réalisé sur un génome de référence du riz (*Oryza sativa japonica* NCBI accession GCA_000005425.2) avec le logiciel d'alignement Bowtie2 (Langmead & Salzberg, 2012). Les algorithmes d'assemblage SPA des (Bankevich et al., 2012) et CAP3 (Huang & Madan, 1999) ont ensuite été utilisés pour réaliser un assemblage *de novo* des reads. Enfin, les similarités nucléotidiques et protéiques entre les contigs obtenus et les reads non assemblés (Brister et al., 2015) ont été recherchées dans les bases de données virales de Genbank avec respectivement les algorithmes BLASTn et BLASTx (Altschul et al., 1990). Une deuxième recherche de similarités nucléotidiques et protéiques de toutes les séquences identifiées comme virales a ensuite été réalisée sur les bases de données génomiques complètes de Genbank (Brister et al., 2015). La première recherche de similarités permet de réduire le nombre de séquences alignées sur les bases générales en éliminant toutes les séquences non virales. La deuxième recherche de similarités permet de valider les assignations taxonomiques virales en éliminant les faux positifs éventuels du premier alignement.

Pour considérer l'infection d'une plante par un phytovirus, nous avons alors établi deux seuils : le premier basé sur le nombre de reads viraux et le second sur la longueur des contigs viraux. Le premier seuil a été établi en triplant le nombre maximal de reads viraux retrouvés dans les 17 témoins négatifs, chacun constitué d'un bout d'une feuille de canne à sucre issue d'une graine indemne de virus (Fuzz). Ce premier seuil a ainsi été fixé à 141 reads (3 fois 47 reads retrouvés dans un des 17 témoins négatifs de canne à sucre). Le second seuil a été basé sur la présence d'au moins un contig phytoviral supérieur à 500 pb, pour considérer l'infection d'une plante par un phytovirus. Alors que le premier seuil apparaît comme « conservateur » et pouvant générer des « faux-négatifs » ; le second seuil, moins conservateur, peut révéler des « faux-positifs ». Ainsi, même si la synthèse des deux analyses basées sur deux seuils de détection différents permet d'avoir un premier regard sur le virome du riz de l'agrosystème HHRTS, il conviendra de confirmer ces résultats par des techniques classiques de biologie moléculaire (RT-PCR, PCR) pour établir la liste des virus présents sur le site d'étude de manière définitive. Cette confirmation a été réalisée dans mes travaux uniquement sur le fijivirus, par une approche de qPCR (voir section 2 de ce chapitre). Elle n'a pas été réalisée sur les autres phyto-virus putativement détectés et décrits ci-dessous.

2.2.2. Résultats et discussion

Sur la base du seuil minimal établi à 141 reads viraux, notre étude de métagénomique virale a permis d'identifier un virus phytopathogène et plusieurs virus appartenant à des familles virales composées de mycovirus et de phytovirus (*Chrysoviridae*, *Partitiviridae* et *Totiviridae*). De façon inattendue, pour un agrosystème défini comme étant peu soumis aux bioagresseurs des plantes, un phytovirus du riz a été assez largement détecté dans notre étude (8,9% de plants infectés). Il s'agit du Southern rice black-streaked dwarf virus (SRBSDV) qui est un virus à double brin d'ARN du genre *Fijivirus* de la famille *Reoviridae*. Ce virus est connu pour être un virus phytopathogène émergent en Asie du Sud-Est. Sa première observation a été réalisée en 2001 dans une parcelle de Guandong (Chine), puis il a été décrit pour la première fois en 2008 (Zhang et al., 2008). Le virus SRBSDV s'est ensuite rapidement généralisé en Asie du Sud-Est, principalement sur toute la partie sud de la Chine et le nord du Vietnam (Hoang et al., 2011, Lv et al., 2017). L'identification de la présence du SRBSDV dans l'agrosystème HHRTS, réputé pour être un agrosystème centenaire, durable et exempt d'épidémies majeures nous a motivés à mettre en place une étude plus approfondie de la caractérisation de ce virus à l'échelle d'une partie de l'agrosystème HHRTS (bassin versant de la rivière Malizhai). L'objectif est de mieux détailler sa prévalence et surtout d'estimer si les variétés modernes sont plus sensibles à ce virus que les variétés traditionnelles. Les résultats de ces travaux sont présentés dans la section suivante de ce chapitre.

Sur la base du seuil minimal de présence établi à 141 reads viraux, notre étude de métagénomique virale a permis d'identifier la présence, chez cinq plants de riz traditionnels, de mycovirus appartenant à trois familles virales distinctes (*Chrysoviridae*, *Partitiviridae* et *Totiviridae*).

Sur la base du deuxième seuil basé sur la longueur des contigs viraux, 22 plants de riz traditionnels étaient putativement infectés par le SRBSDV (11,6% des plants de riz analysés). Par ailleurs, 19 plants de riz traditionnels étaient putativement infectés par des mycovirus appartenant à trois familles virales distinctes (*Endornaviridae*, *Partitiviridae* et *Totiviridae*). Il est à noter que l'inventaire des phytovirus par l'utilisation des deux seuils est légèrement différent concernant les mycovirus.

Le seuil « reads » a détecté un chrysovirus dans une plante non détectée par le seuil « contig » cependant le seuil contig a permis de détecter un endornavirus probablement associé à un champignon (34% d'identité protéique avec *Ceratobasidium endornavirus A*) dans une autre

plante passé inaperçue avec le seuil « reads ». Cette observation suggère que l'inventaire des phytovirus par les deux seuils reste incomplet et que l'utilisation de techniques plus sensibles et plus spécifiques de détection devrait être utilisée pour confirmer les analyses de métagénomique virale. Il est possible que la présence de certains de ces mycovirus puisse avoir un rôle de biocontrôle des champignons pathogènes du riz et ainsi contribuer au maintien d'un faible niveau de maladies fongiques observées dans l'agrosystème HHRTS. Il a par ailleurs, déjà été montré l'existence d'activités de biocontrôle de virus appartenant aux familles *Partitiviridae* et *Totiviridae* sur des maladies fongiques des plantes (de la Paz Giménez-Pecci et al., 2002, Yadav et al., 2015). Il a notamment été montré que des virus de la famille *Totiviridae* peuvent avoir un effet de réduction de la virulence de *Pyricularia oryzae* (Moriyama et al., 2018, Urayama et al., 2012), qui est considéré comme le pathogène fongique majeur du riz dans le monde (Skamnioti & Gurr, 2009). Ces observations pourraient expliquer en partie l'absence d'épidémies majeures de pyriculariose dans l'agrosystème HHRTS. Elles nécessitent d'être confrontées avec des tests de détection des isolats de mycovirus sur les souches de *Pyricularia oryzae* isolées à partir d'échantillons de l'agrosystème HHRTS et de réaliser des tests sur leur pouvoir de biocontrôle potentiel de *Pyricularia oryzae* sur riz.

Sur la base du deuxième seuil reposant sur la longueur des contigs viraux, deux phytovirus ont aussi putativement été détectés. Il s'agit d'un polerovirus et d'un virus appartenant à la famille *Tombusviridae*, chacun infectant une seule plante. Un contig de 875 pb a été assemblé pour ce dernier virus qui présente une similarité nucléotidique de 71,7% avec le Maize chlorotic mottle virus. Des travaux de caractérisation complémentaires seraient nécessaires pour préciser si ce virus est un variant très divergent du machlomovirus du maïs ou s'il représente une nouvelle espèce virale. Concernant le polerovirus, trois contigs de 1363 pb, 1382 pb et 2041 pb ayant respectivement des identités nucléotidiques de 90,9%, 86,7% et 86,4% avec *Panicum distortion mosaic virus* (PDMV) ont été assemblés. Il est donc tout à fait probable que le virus retrouvé dans seulement un échantillon de la variété de riz traditionnelle appartenant à la sous-espèce de riz *Oryza sativa japonica* ait été identifié comme étant un isolat de PDMV. Ce virus a été initialement isolé d'un plant de millet commun en Corée du Sud. Son génome complet a été déposé sur NCBI GenBank sous le numéro d'accès LC424839, mais n'a pas fait encore l'objet d'une publication. Ce virus, dont le génome a déjà été caractérisé, et n'étant présent que dans un seul échantillon de notre étude n'a pas fait l'objet d'une attention supplémentaire dans ce travail de thèse.

2.2.3. Conclusion

Ce travail initial de métagénomique virale a permis de mettre en évidence la présence du SRBSDV, virus pathogène du riz, dans les rizières du village Malizhai. Les deux seuils de détection que nous avons utilisés, permettent d'estimer une prévalence de l'ordre de 8,9-11,6%, suggérant qu'une épidémie de SRBSDV était en train de se dérouler lorsque nous avons réalisé notre échantillonnage. Comme mentionné précédemment, ce résultat nous a encouragé à caractériser plus finement l'épidémiologie de ce virus dans l'agrosystème HHRTS. Cette étude sera présentée dans la section suivante de ce chapitre. Par ailleurs, ces analyses de métagénomique virale ont révélé que différents mycovirus étaient présents parmi plusieurs plants de riz traditionnel et étaient absents des feuilles de riz moderne. Il serait intéressant de tester l'hypothèse selon laquelle ces virus pourraient avoir un rôle de biocontrôle des champignons phytopathogènes au sein de l'agrosystème HHRTS. Il serait aussi intéressant de mieux comprendre si ces virus sont significativement plus présents chez les variétés traditionnelles. Dans le cas où les deux hypothèses précédentes seraient vérifiées, le remplacement des variétés traditionnelles par des variétés modernes participerait probablement à une réduction du biocontrôle des maladies fongiques dans le système des terrasses du Yuanyang. Enfin, nos travaux de métagénomique ont révélé la présence de deux autres virus phytopathogènes, chacun présent sur seulement un plant de riz. Une veille épidémiologique de ces deux virus serait importante à mettre en place pour voir si leur prévalence respective reste faible dans le temps et dans l'espace ou si un phénomène d'émergence peut se produire dans les années à venir.

2.3. Section 2: Emergence of Southern Rice Black-Streaked Dwarf Virus in the Centuries-Old Chinese Yuanyang Agrosystem of Rice Landraces

Article

Emergence of Southern Rice Black-Streaked Dwarf Virus in the Centuries-Old Chinese Yuanyang Agrosystem of Rice Landraces

Pascal Alonso ^{1,2} , Pierre Gladieux ^{2,3}, Oumaima Moubset ^{1,2}, Pei-Jung Shih ^{1,2,4} , Pierre Mournet ^{5,6} , Julien Frouin ^{5,6}, Laurence Blondin ^{1,2} , Romain Ferdinand ^{1,2}, Emmanuel Fernandez ^{1,2}, Charlotte Julian ^{1,2}, Denis Filloux ^{1,2}, Henry Adreit ^{1,2}, Elisabeth Fournier ^{2,3}, Aurélie Ducasse ^{2,3}, Vladimir Grosbois ⁷, Jean-Benoit Morel ^{2,3}, Huichuan Huang ⁸ , Baihui Jin ⁸, Xiahong He ^{8,9}, Darren P. Martin ¹⁰, Christian Vernière ^{1,2}  and Philippe Roumagnac ^{1,2,*}

- ¹ CIRAD, BGPI, 34398 Montpellier, France; pascal.alonso@cirad.fr (P.A.); oumaima.moubset@etu.umontpellier.fr (O.M.); y199404_20@yahoo.com.tw (P.-J.S.); laurence.blondin@cirad.fr (L.B.); romain.ferdinand@cirad.fr (R.F.); emmanuel.fernandez@cirad.fr (E.F.); charlotte.julian@cirad.fr (C.J.); denis.filloux@cirad.fr (D.F.); henri.adreit@cirad.fr (H.A.); christian.vernieres@cirad.fr (C.V.)
- ² BGPI, INRA, CIRAD, SupAgro, Univ Montpellier, 34398 Montpellier, France; pierre.gladieux@inra.fr (P.G.); elisabeth.fournier@inra.fr (E.F.); aurelie.ducasse@inra.fr (A.D.); jean-benoit.morel@cirad.fr (J.-B.M.)
- ³ INRA, BGPI, 34398 Montpellier, France
- ⁴ Department of Plant Pathology, National Chung Hsing University, Taichung 402, Taiwan
- ⁵ CIRAD, UMR AGAP, 34398 Montpellier, France; pierre.mournet@cirad.fr (P.M.); julien.frouin@cirad.fr (J.F.)
- ⁶ AGAP, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, 34398 Montpellier, France
- ⁷ CIRAD, UMR ASTRE, 34398 Montpellier, France; vladimir.grosbois@cirad.fr
- ⁸ State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan Agricultural University, Kunming 650201, China; absklhch@gmail.com (H.H.); 18214562004@163.com (B.J.); hexiahong@hotmail.com (X.H.)
- ⁹ Southwest Forestry University, Kunming 650224, China
- ¹⁰ Computational Biology Group, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town 4579, South Africa; darrenpatrickmartin@gmail.com
- * Correspondence: philippe.roumagnac@cirad.fr; Tel.: +33(0)-4-99-62-48-53; Fax: +33(0)-4-99-62-48-48

Received: 23 September 2019; Accepted: 15 October 2019; Published: 25 October 2019



Abstract: Southern rice black-streaked dwarf virus (SRBSDV), which causes severe disease symptoms in rice (*Oriza sativa* L.) has been emerging in the last decade throughout northern Vietnam, southern Japan and southern, central and eastern China. Here we attempt to quantify the prevalence of SRBSDV in the Honghe Hani rice terraces system (HHRTS)—a Chinese 1300-year-old traditional rice production system. We first confirm that genetically diverse rice varieties are still being cultivated in the HHRTS and categorize these varieties into three main genetic clusters, including the modern hybrid varieties group (MH), the Hongyang improved modern variety group (HY) and the traditional *indica* landraces group (TIL). We also show over a 2-year period that SRBSDV remains prevalent in the HHRTS (20.1% prevalence) and that both the TIL (17.9% prevalence) and the MH varieties (5.1% prevalence) were less affected by SRBSDV than were the HY varieties (30.2% prevalence). Collectively we suggest that SRBSDV isolates are freely moving within the HHRTS and that TIL, HY and MH rice genetic clusters are not being preferentially infected by particular SRBSDV lineages. Given that SRBSDV can cause 30–50% rice yield losses, our study emphasizes both the need to better monitor the disease in the HHRTS, and the need to start considering ways to reduce its burden on rice production.

Keywords: Southern rice black-streaked dwarf virus; rice; Honghe Hani rice terraces system; host genotyping; virus prevalence; virus emergence

1. Introduction

Southern rice black-streaked dwarf virus (SRBSDV) causes the emergent rice black-streaked dwarf disease that has for the last decade been threatening rice-growing areas of northern Vietnam, southern Japan and southern, central and eastern China [1–4]. While the first report of black-streaked dwarf disease was in 2001 from a Chinese rice field in Guangdong, SRBSDV was only shown to be a novel fijivirus in 2008 [5,6]. Symptoms of SRBSDV infections include severe stunting, darkening of leaves, and white waxy or black-streaked swellings along stem veins that can lead to 30–50% rice yield losses [1,2]. The virus is transmitted by the white-backed planthopper *Sogatella furcifera*, but is not seed transmissible [2]. Besides rice, SRBSDV is able to infect several other species, including maize, sorghum and various uncultivated grasses [5,6]. The genome of SRBSDV is composed of ten linear dsRNA segments between 1.8 and 4.5 kb in length which encode twelve proteins [7]. Observed SRBSDV genetic diversity is low: for example, all characterized segment 8 sequences of Chinese SRBSDV isolates share more than 93.9% nucleotide sequence identity [8]. This low diversity has been interpreted as indirect evidence that SRBSDV has only recently emerged.

While SRBSDV is now known to be present in all provinces of Southern China, its prevalence and geographical distribution within each of these provinces remains unknown. Additionally, given that the prevalence of other plant viruses has been found to be higher in cultivated areas than in natural settings [9], it would be interesting to determine whether the relative prevalence of SRBSDV is lower in rice agronomic systems that use mixtures of landraces in comparison to those that use genetically-uniform modern varieties. For example, the prevalence and distribution of SRBSDV in the Honghe Hani rice terraces system (HHRTS) in the Yunnan province—a 1300-year-old traditional rice production system [10]—has never been precisely studied.

The HHRTS was developed over centuries by the Hani people and, amongst other peculiarities, is characterized by the cultivation of a remarkably diverse variety of locally bred rice landraces [10]. In 2008, each village in this mountainous region grew complex heterogeneous mosaics of rice genotypes drawn from at least 47 diverse rice landraces [10]. The Hani people have created stretched terraces that cascade down hillside slopes from 800 to 2000 m above sea level. Each terrace is composed of thin stretched fields within which only one rice landrace is grown. The diversity of landraces used in the HHRTS has for centuries apparently minimized the impacts of both biotic and abiotic stressors on rice production in this region [10]. For example, the long term cultivation of genetically diverse HHRTS landraces may have led today to the situation where two different lineages of the rice blast fungus, *Pyricularia oryzae*, are each locally adapted to either the *indica* or the *japonica* rice landraces that are grown in the terraces [11]. This specialization of different *P. oryzae* populations may limit the spread of this pathogen in the HHRTS, which, by extension, has reduced the impact of rice blast on rice yields in the terraces [11]. Given the relatively light burden of bacterial and fungal diseases affecting HHRTS landraces, in 2010, the Food and Agriculture Organization (FAO) designated the HHRTS as a “Globally Important Agricultural Heritage System” and, in 2013, the United Nations Educational, Scientific, and Cultural Organization (UNESCO) listed it as a World Cultural Heritage site.

However, the social, economic and environmental transformations that are presently occurring in the HHRTS are threatening the existence of this unique, highly successful, sustainable and durable agrosystem [10]. Specifically, chemical fertilizers are being increasingly used and genetically divergent landraces are being gradually replaced by high-yielding modern hybrid varieties [12]. A recent survey involving interviews with HHRTS farmers has revealed that those farmers who have attempted to incorporate modern fertilizers and hybrid varieties to achieve higher yields, have noticed a clear increase in the prevalence and severity of rice diseases [13].

Here we attempt to both quantify the prevalence of SRBSDV in the HHRTS and determine whether this emergent virus has a greater impact on modern high yielding rice varieties than it does

on traditional HHRTS landraces. We first confirm that genetically diverse rice varieties are still being cultivated in the HHRTS and categorize these varieties into three main genetic clusters: the traditional *indica* landraces group (TIL); the Hongyang group (HY) that contains improved modern varieties sharing recent ancestry with the TIL varieties; and the modern hybrid varieties group (MH) that includes commercial varieties usually grown outside the HHRTS. We show over a period of two years that SRBSDV remains highly prevalent in the HHRTS (20.1% prevalence) and that both the TIL (17.9% prevalence) and the MH varieties (5.1% prevalence) were less affected by SRBSDV than were the HY varieties (30.2% prevalence)

2. Material and Methods

2.1. Study Area

While the upper reaches of the Malizhai River Basin in the vicinity of Xinjie town in the Honghe Hani Yi Autonomous Prefecture (Yunnan province, China) is representative of the cultural landscape of the Hani rice terraces [10], we focused our study on villages located in this river basin. The Hani people are perpetuating land-management practices related to the cultivation of irrigated rice terraces that include profound and long-standing understanding of environmental conditions such as landforms, soil conditions, vegetation types, and hydrology. However, rice landraces are being gradually replaced by high-yielding modern hybrid varieties in the Malizhai River Basin [12] and part of Hani farmers from the Malizhai River Basin are adopting mixed landrace/modern variety systems [13]. Specifically, Hani farmers are transplanting landrace and modern varieties seedlings into the fields after 60- and 40-days germination in the nursery plot, respectively. Neither chemical fertilizer nor fungicide and pesticide are usually applied in the field during the whole growth stages of the plants.

2.2. Plant Sampling

Seven villages from the upper Malizhai River Basin were prospected, including one village (Malizhai altitude of 1570–1608 m) in 2016 and six villages (Xiaoshuijing, 1781 m alt.; Tuguozhai, 1778–1793 m alt.; Shuibulong, 1647–1666 m alt.; Huangcaolin, 1732–1747 m alt.; Gingko, 1629 m alt.; and Dayutang, 1769–1805 m alt.) in 2018 (Figure S1). Both traditional landraces and modern rice varieties were cultivated in these seven villages. In each village, approximately equal numbers of leaf samples from modern varieties and traditional landraces were collected by farmers and government technicians without regard for the presence of disease symptoms. Seven to ten plants were collected in each of the sampled rice fields (Table 1). Collectively, these accounted for 186 samples collected in 2016 and 173 samples collected in 2018 (Table 1). All of these plants were assumed to have grown under the same weather, water, soil and nutrient conditions.

Table 1. Genotyping of rice varieties from the Yuanyang rice terraces of China and screening of these varieties for the presence of Southern rice black-streaked dwarf virus using quantitative real-time PCR (qPCR) assays. HY: Hongyang improved modern variety group, MH: modern hybrid variety group, SRBSDV: Southern rice black-streaked dwarf virus, TIL: traditional *indica* landraces group and Na: not assigned.

Sampling Field	Variety Group	Year	Rice Variety Name	Village	Number of Samples	SRBSDV Positive Samples	SRBSDV Negative Samples
YYT1	HY	2016	Chepugu	Malizhai	9	5	4
YYT6	HY	2016	Chepugu	Malizhai	9	4	5
YYT10	HY	2016	Jinpinggu	Malizhai	10	4	6
Plot_1	HY	2018	Hongyang	Gingko	10	1	9
Plot_3	HY	2018	Hongyang	Gingko	10	3	7
Plot_8	HY	2018	Hongyang	Huangcaolin	9	2	7
Plot_13	HY	2018	Hongyang	Xiaoshuijing	10	0	10
Plot_16	HY	2018	Hongyang	Tuguozhai	10	4	6
Plot_19	HY	2018	Hongyang	Shuibulong	9	0	9
Plot_22	HY	2018	Hongyang	Dayutang	10	6	4
YYM1	MH	2016	Mingliangyou 527	Malizhai	10	1	9
YYM3	MH	2016	Mingliangyou 528	Malizhai	8	1	7
YYM4	MH	2016	Hefeng 177	Malizhai	10	2	8
YYM5	MH	2016	Zhongyou 177	Malizhai	10	0	10
YYM6	MH	2016	Liangyou 2186	Malizhai	10	0	10
YYM7	MH	2016	Guofeng 1	Malizhai	10	0	10
YYM8	MH	2016	Liangyou 725	Malizhai	10	0	10
YYM9	MH	2016	Liangyou 2161	Malizhai	10	0	10
YYT2	TIL	2016	Lubaigu	Malizhai	10	1	9
YYT3	TIL	2016	Zaogu	Malizhai	10	0	10
YYT5	TIL	2016	Epugu	Malizhai	10	1	9
YYT9	TIL	2016	Nuogu	Malizhai	10	4	6
YYT11	TIL	2016	Luhonggu	Malizhai	10	0	10
Plot_5	TIL	2018	Acuce	Gingko	10	2	8
Plot_6	TIL	2018	Acuce	Gingko	10	3	7
Plot_7	TIL	2018	Acuce	Huangcaolin	9	1	8
Plot_9	TIL	2018	Acuce	Huangcaolin	10	5	5
Plot_10	TIL	2018	Acuce	Xiaoshuijing	10	0	10
Plot_11	TIL	2018	Acuce	Xiaoshuijing	7	3	4
Plot_15	TIL	2018	Acuce	Tuguozhai	10	0	10
Plot_18	TIL	2018	Acuce	Shuibulong	9	0	9
Plot_20	TIL	2018	Acuce	Shuibulong	10	2	8
Plot_23	TIL	2018	Acuce	Dayutang	10	4	6
YYT4	Na (outlier)	2016	Nuogu	Malizhai	10	0	10
YYT7	Na (outlier)	2016	Jianshuigu	Malizhai	10	0	10
YYT8	japonica	2016	Honglueduolu	Malizhai	10	0	10
Plot_12	japonica	2018	Unknown	Xiaoshuijing	10	1	9

2.3. Genotyping-by-Sequencing of Rice Landraces

Total genomic DNA extractions were performed from 30 mg of dried individual leaf samples. Plant cell were lysed using the mixed alkyltrimethylammonium bromide (MATAB) buffer [14] at 72 °C for 30 min. DNA was then extracted using an automated method on a Biomek FXP instrument (Beckman Coulter, CA, USA) [15] that uses the NucleoMag Plant Kit (Macherey–Nagel, Germany). DNA concentrations were quantified with Hoescht dye and a Fluoroskan Ascent FL fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). Genomic DNA quality was further checked using agarose gel electrophoresis. Two genomic libraries (one library containing the 186 samples collected in 2016 and the other the 173 samples from 2018) were prepared as previously described [16] using *ApeKI* restriction enzyme digestion (New England Biolabs, Hitchin, UK) from 100 ng of DNA per sample. A ligation reaction was then carried out using T4 DNA ligase (New England Biolabs, Hitchin, UK) and barcoded adapters at 22 °C for 30 min [16], then ligase was inactivated by heating at 65 °C for 20 min. Barcoded samples were then pooled and PCR-amplified. PCR reactions were carried out in a 50 µL volume consisting of 5 µL of pooled and barcoded samples, 2 µL of each primer (10 µmol/µL), 25 µL of NEB 2X *Taq* Master Mix (NEB#M0270S; New England Biolabs, Hitchin, UK) and RNase-free water to 50 µL. The PCR program was as follows: 72 °C for 5 min, 98 °C for 30 s, 18 cycles at 98 °C for 10 s, 65 °C for 30 s and 72 °C for 30 s with a final 72 °C extension for 5 min. The PCR-amplified

libraries were purified using the Wizard PCR prep DNA purification system from Promega (Madison, WI, USA) and verified with the Agilent D5000 ScreenTape instrument (Santa Clara, CA, USA). While the 2016-library was sequenced at the GeT-PlaGe platform at Toulouse (France) using a single lane on an Illumina HiSeq system (2 × 150 bp sequencing), the 2018 library was sequenced by Genewiz (Leipzig, Germany) using a single lane on an Illumina HiSeq system (2 × 150 bp sequencing). The raw sequence read data from these two Illumina HiSeq runs are available from the NCBI Sequence Read Archive, under the study accession number: PRJNA573048

2.4. Analysis of Rice Genotyping-by-Sequencing Data

Read processing, read mapping and single nucleotide polymorphism (SNP) calling were carried out using bioinformatic pipelines designed using Toggle (<https://github.com/SouthGreenPlatform/TOGGLE>). Reads were aligned to the Nipponbare rice reference genome (https://plants.ensembl.org/Oryza_sativa/Info/Index) using Burrows Wheeler alignment (BWA) [17], and SNPs were called using UnifiedGenotyper in GATK v. 3.8 [18]. Only positions in coding sequences, with DP > 3 and MQ > 20, and with less than 50% missing data were retained. The final dataset used for genotyping consisted of 12,112 SNPs. Accessions Plot-22-P3 and YYT10-A were removed from the data as proportions of missing data in these exceeded 95%.

We constructed a neighbor-network using SplitsTree 4.13 [19], to visualize evolutionary relationships between the *indica* rice genotypes while taking the possibility of recombination or incomplete lineage sorting into account. We also used the program sNMF (<http://membres-timc.imag.fr/Olivier.Francois/snmf/index.htm>) to infer population subdivision by partitioning genotypes into K ancestral populations and estimating individual ancestry coefficients in the K populations. We ran sNMF for K ranging from 1 to 20, and for each K value 10 replications were performed. We used the Greedy algorithm in CLUMPP 1.1.2 (<http://web.stanford.edu/group/rosenberglab/clumpp.html>) to identify runs belonging to the same mode (i.e., representing the same clustering solution) and we randomly selected one representative of the major mode for graphical representation as a stacked barplot using the Matplotlib package in Python.

2.5. RNA Extraction and Detection of SRBSDV by Real-Time Quantitative PCR

Virions were initially concentrated from approximately 100 mg of rice leaf tissue using the virion-associated nucleic acids (VANA) protocol as previously described [20]. Total RNA of the virion preparation was extracted with the QIAGEN® RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) as described by the manufacturer. Three-hundred and fifty-nine plant sample RNAs (186 from 2016 and 173 from 2018) were tested for the presence of SRBSDV using quantitative real-time PCR (qPCR). The previously designed primers [21], s9-1-F (5'-AACGACCAACCAACAAGA-3') and s9-1-R (5'-GTTCCATCAATGAGGTAGTTC-3') were used. The reverse transcription and the qPCR reaction were performed using the QIAGEN® OneStep RT-PCR Kit and the Invitrogen™ Quant-iT™ PicoGreen™ dsDNA reagent on a Stratagene Mx3000P instrument. The 25 µL RT-qPCR reaction mix consisted of 2 µL of eluted RNA, 13 µL of RNase-free water, 5 µL of QIAGEN OneStep RT-PCR buffer (5×), 1 µL of dNTP mix (10 mM), 1.25 µL of each primer (10 µM), 1 µL of PicoGreen 1× and 1 µL of QIAGEN OneStep RT-PCR enzyme mix. The RT-qPCR program was as follows: 50 °C for 30 min, 95 °C for 15 min, 40 cycles at 94 °C for 1 min, 58 °C for 1 min and 72 °C for 1 min with a final 72 °C extension for 10 min followed by a melt curve step (from 55 °C, gradually increasing 0.5 °C/s to 95 °C, with acquisition data every 1 s).

2.6. Reverse-Transcription PCR, Partial SRBSDV Segment 8 Sequencing and Sequence Analysis

A fragment of SRBSDV segment 8 encoding a putative core structural protein [21] was amplified from 26 SRBSDV isolates sampled from the HHRTS (including 20 isolates from 2016 and six isolates from 2018) by RT-PCR using primers SRBSDV_S8_F3 (5'-GGG TTG ATT CCT TTG GT-3') and SRBSDV_S8_R3 (5'-ACG GGA TTG TCT CCT TTG-3') [8]. The RT-PCR reaction was performed as described above

without adding PicoGreen reagent. The PCR product was analyzed by electrophoresis on a 1.0% agarose gel in TAE 1X buffer stained with ethidium bromide and visualized under UV light. The SRBSDV amplification product had an expected size of 776 bp. Amplicons were directly sequenced using automated Sanger sequencing (Genewiz, South Plainfield, NJ, USA). SRBSDV segment 8 fragments were further aligned using MUSCLE [22] as implemented in MEGA with default settings [23] and manually trimmed (down to 624 nt in length). Pairwise identity analyses of partial segment 8 sequences were carried using SDT v1.2 [24].

2.7. Phylogenetic Analysis

Twenty-six SRBSDV segment 8 fragments (624 nt in length) and all 29 of the SRBSDV segment 8 sequences that were publicly available in June 2019 were aligned using MUSCLE [22] as implemented in MEGA (with default settings). A maximum likelihood phylogenetic tree was constructed using PhyML3 [25] with the T92 nucleotide substitution model (selected as best fit by MEGA) and 1000 bootstrap replicates were used to test the support of branches.

2.8. Statistical Analyses

We applied a generalized linear model (GLM) implemented in *R Base* (<http://www.r-project.org>) to analyze proportions of SRBSDV-infected plants (number of diseased plants per plot / total number of collected plants in each plot) in 2016 and 2018. Sampling plots were split into three rice variety groups: the TIL group, the HY improved modern variety group and the MH variety group (Table 1). As variation in the number of SRBSDV-infected plants given the number of plants collected was greater than that expected under standard binomial models, a phenomenon referred to as overdispersion, the beta-binomial model was further fitted to our data. The effect of rice variety group and the year of sampling were tested using two GLM models: (1) a model where the two explanatory variables were considered as having additive effects (variety group + year of sampling); and (2) a model additionally including an interaction between the two explanatory variables (variety group + year of sampling + variety group: year of sampling). Tukey's multiple comparison test (function Tukey HSD in *R*) was further used to test for pairwise differences in SRBSDV prevalence among variety groups.

3. Results and Discussion

3.1. Highly Genetically Diverse TIL, HY and MH Varieties Are Cultivated in the HHRTS

Rice genotyping-by-sequencing confirmed that the HHRTS hosts a large diversity of rice varieties, including traditional *indica* or *japonica* local rice subspecies and modern *indica* varieties that are widely grown in other regions of China and elsewhere in South-East Asia (Table 1). The genotyping also confirmed that only one rice variety was grown in each of the small rice fields that were sampled in this study (Figure 1A and Figure S1). Specifically, we show that, besides TILs, two groups of modern *indica* varieties were being grown in the HHRTS in 2016 and 2018: (1) the HY variety group; and (2) the MH variety group (Figure 1A). The admixture analysis revealed that the HY varieties that have been in use in the HHRTS since 2010 [13] share recent ancestry with the TIL varieties ($K = 2$ to $K = 5$ models in Figure 1B).

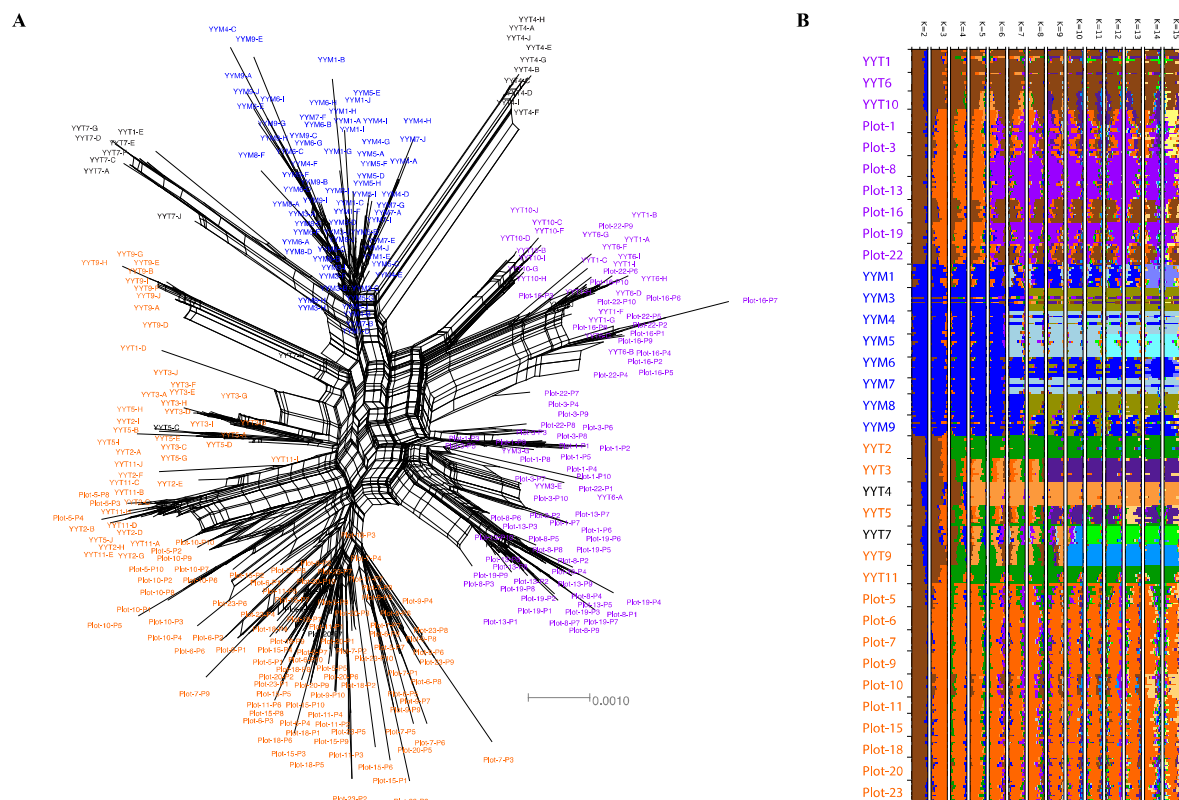


Figure 1. (A) Neighbor-Net split decomposition network indicating the relationships between rice accessions based on 12,112 analyzed SNPs. Plant samples assigned to the HY, MH and TIL *indica* variety groups are labelled in purple, blue and orange, respectively. Plant samples that were not assigned to one of the three variety groups (“outlier varieties”) are labelled in black. Characters A to J or P1 to P10 (e.g., YYT5-J or Plot-23-P2) refer to individual plants that were collected from each rice field (7–10 plants were collected in each field, Table 1) (B) Ancestry proportions within accessions for sNMF models [26] from K = 2 to K = 15 ancestral populations. Each horizontal bar represents the proportion of ancestry within a single accession, with colors corresponding to ancestral populations.

Accordingly, the branches connecting the HY and TIL varieties in a split decomposition network constructed from 12,112 SNPs were more reticulated than the branches connecting HY or TIL varieties to the MH varieties, indicating conflicting phylogenetic signals caused by recent shared ancestry between HY and TIL, but not between MH and HY/TIL (Figure 1A). This split decomposition network structure suggests that the MH varieties that are commercially grown outside the HHRTS have a narrow genetic base and have likely been bred from rice lineages which do not share any recent ancestries with either TIL or HY local varieties. The lack of shared ancestry between MH and TIL/HY in our admixture analysis also supports this hypothesis (Figure 1B). Finally, two varieties (YYT4 and YYT7, Table 1) that were considered to be traditional landraces by the farmers, did not cluster phylogenetically with the TIL, HY or MH groups (Figure 1A) and are therefore hereafter referred to as “outlier” varieties. Collectively these results demonstrate the presence of strong rice population structure within the HHRTS with a group of *japonica* varieties, three large groups of *indica* varieties and two other *indica* varieties, reflecting divergent breeding histories.

3.2. SRBSDV Prevalence and Distribution

A total of 359 rice leaf samples from the HHRTS (186 from 2016 and 173 from 2018) were tested for the presence of SRBSDV, 60 of which (20.1%) tested positive by qPCR (Table 1). These consisted of 23 samples collected in 2016 (12.4% of the 2016 samples, Table 1) and 37 collected in 2018 (21.4% of the 2018 samples, Table 1). SRBSDV was detected in every village where plants were sampled, and virus

prevalence ranged from 7.1% in Shuibulong village to 50% in Dayutang village (Table 2). These results indicate that, as is the case elsewhere in South China and northern Vietnam [1–4], SRBSDV is highly prevalent in the HHRTS. This brings into question the assumption that, in general, viruses causing rice diseases have low prevalence in the HHRTS [10]. Although this indicates that SRBSDV may at present be a major biotic constraint on rice production in the HHRTS, it is plausible that, given enough time, the HHRTS farmers will eventually select additional SRBSDV-tolerant/resistant varieties for use in the terraces.

Table 2. Screening of rice samples from the Yuanyang rice terraces of China for the presence of Southern rice black-streaked dwarf virus using qPCR assays. Rice varieties are grouped in rice variety groups (rows) or villages from which they were collected (columns). HY: Hongyang improved modern variety group, MH: modern hybrid variety group, TIL: traditional *indica* landraces group and Na: not assigned. SRBSDV prevalence is indicated within brackets and bold characters.

Year	2016			2018			
Village	Malizhai (12.4%)	Xiaoshuijing (10.8%)	Tuguozhai (20%)	Shuibulong (7.1%)	Huangcaolin (28.6%)	Gingko (22.5%)	Dayutang (50%)
TIL (17.9%)	6/50 (12%)	3/17 (17.6%)	0/10	2/19 (10.5%)	6/19 (31.6%)	5/20 (25%)	4/10 (40%)
HY (30.2%)	13/28 (46.4%)	0/10	4/10 (40%)	0/9	2/9 (22.2%)	4/20 (20%)	6/10 (60%)
MH (5.1%)	4/78 (5.1%)						
Na (outlier)	0/20						
<i>Japonica</i> (5.3%)	0/10	1/10 (11.1%)					

While long-term monitoring of SRBSDV prevalence in the HHRTS will be needed to determine how successfully such selective breeding efforts will be, a reasonable first step towards controlling the virus in the short term will be to determine which of the HY, MH and TIL varieties are most susceptible to infection by SRBSDV. We found that among the rice plants that were detectably infected with the virus, the HY varieties have the highest SRBSDV prevalence (30.2%) with infected HY plants being found in 8/10 sampled HY fields (Tables 1 and 2). Both the TIL (17.9% prevalence occurring in 10/15 TIL fields) and MH varieties (5.1% prevalence in 3/8 MH fields) were less affected by SRBSDV than were the HY varieties (Tables 1 and 2). Furthermore, none of the 20 samples from the two “outlier” *indica* landraces, which are phylogenetically most closely related to the MH varieties (Figure 1A), were detectably infected by SRBSDV (Table 1). Finally, the *japonica* plant samples were rarely affected by SRBSDV (5% prevalence in 1/2 *japonica* fields, Table 1).

Given the observed differences in the detected prevalence of SRBSDV in the different rice variety groups, we tested whether the variety groups (HY, TIL and MH) or the year of sampling (2016, 2018) displayed statistically significant differences in virus prevalence using a GLM-based approach. The model including the interaction between variety group and year of sampling revealed that there was no significant interaction between these explanatory variables ($p = 0.067$). Consequently, we selected the additive model that indicated that the SRBSDV prevalence observed for the different variety groups were significantly different ($p = 0.012$). In addition, pairwise comparisons using the Tukey HSD test indicated that there was significantly higher prevalence of the virus in the HY varieties than in the MH varieties ($p = 0.024$) but no significant difference in prevalence between the HY and TIL varieties ($p = 0.241$), or between the TIL and MH varieties ($p = 0.160$).

Interestingly, there was no significant difference in SRBSDV prevalence between 2016 and 2018 ($p = 0.710$). The sustained prevalence of the virus between 2016 and 2018 suggests that SRBSDV is possibly now endemic in the HHRTS system.

3.3. SRBSDV Genetic Diversity

The degree of genetic diversity within the amplified HHRTS SRBSDV genome segment 8 sequences was compared to that of 29 other publicly available SRBSDV segment 8 sequences collected since 2003 from several regions of China and Vietnam. The 26 fragments (624 bp long after alignment and trimming of the 5'- and 3'-ends of the sequences; GenBank accession numbers: MN244953- MN244978) from the HHRTS share between 98.2% and 100% pairwise identity with one another and between 94.1% and 100% pairwise identity with the 29 publicly available SRBSDV segment 8 sequences. Interestingly, all SRBSDV sequences from rice plants are genetically highly homogeneous (sharing between 98.2% and 100% pairwise identity with one another) whereas they share slightly lower identities with SRBSDV isolates sampled from maize (between 94.1% and 99% pairwise identity).

Overall, these degrees of sequence identity are relatively high, and are consistent with SRBSDV having emerged and spread to rice plants throughout Southern China and Northern Vietnam in the recent past. Specifically, the genetic distance between the two most divergent Chinese and Vietnamese rice isolates is only 0.018 nucleotide differences per site implying that the genetic distance of these most divergent isolates to their most recent common ancestor is approximately $0.018/2 = 0.009$ nucleotide differences per site. If SRBSDV did indeed emerge in approximately 2000 this would suggest a nucleotide substitution rate in the order of approximately 5×10^{-4} substitutions per site per year: a substitution rate which is in the middle of the range of those estimated previously for other double stranded RNA viruses [27–32].

A maximum likelihood phylogenetic tree containing all 55 of the available SRBSDV partial segment 8 sequences together with homologous genome fragments of three other closely related fijivirus species (included for rooting purposes) indicated that SRBSDV isolates from the Chinese Hainan region and the SRBSDV clade (including all isolates from all other regions) are sister groups (Figure 2).

It is noteworthy that, while all of these isolates from the Hainan region were isolated from maize, all but one of the other isolates from the SRBSDV clade were isolated from rice plants collected from China and Vietnam (Figure 2). These results are consistent with SRBSDV having originated in maize—possibly in the Hainan region—prior to it spreading into the rice fields of South China and Northern Vietnam with the recurrent annual migrations of viruliferous white back planthopper *Sogatella furcifera*. However, properly testing this hypothesis will require both the genomic sequence characterization and analysis of many more SRBSDV isolates from rice and maize plants throughout the entire current geographical range of the virus, and an accurate estimate of the SRBSDV nucleotide substitution rate. With this data in hand it should be possible to determine the most likely time when, and place where, the SRBSDV isolates that are responsible for rice black-streaked dwarf disease first arose.

Although all of the SRBSDV isolates from the HHRTS clustered within the SRBSDV “rice clade” (Figure 2) there was no evidence to indicate that they may have all shared a more recent common ancestor with one another than with other SRBSDV isolates from China and Vietnam. This suggests that there have likely been multiple (and possibly ongoing) transmissions of SRBSDV into the HHRTS.

Also, within the HHRTS we were unable to find any phylogenetic evidence to support the hypothesis that particular rice variety groups (i.e., TIL, HY and MH) were being preferentially infected by particular SRBSDV lineages. Rather, the SRBSDV isolates obtained from the three variety groups, or from different villages in different years, were interspersed throughout the maximum likelihood tree: a pattern indicating both the persistence of viruses between years, and frequent transmissions of viruses between the variety groups and between the villages (Figure 2). Collectively the phylogenetic evidence supports the hypotheses that SRBSDV isolates are freely moving within the HHRTS and that no obvious rice genotype-specific specialization of SRBSDV isolates has occurred.

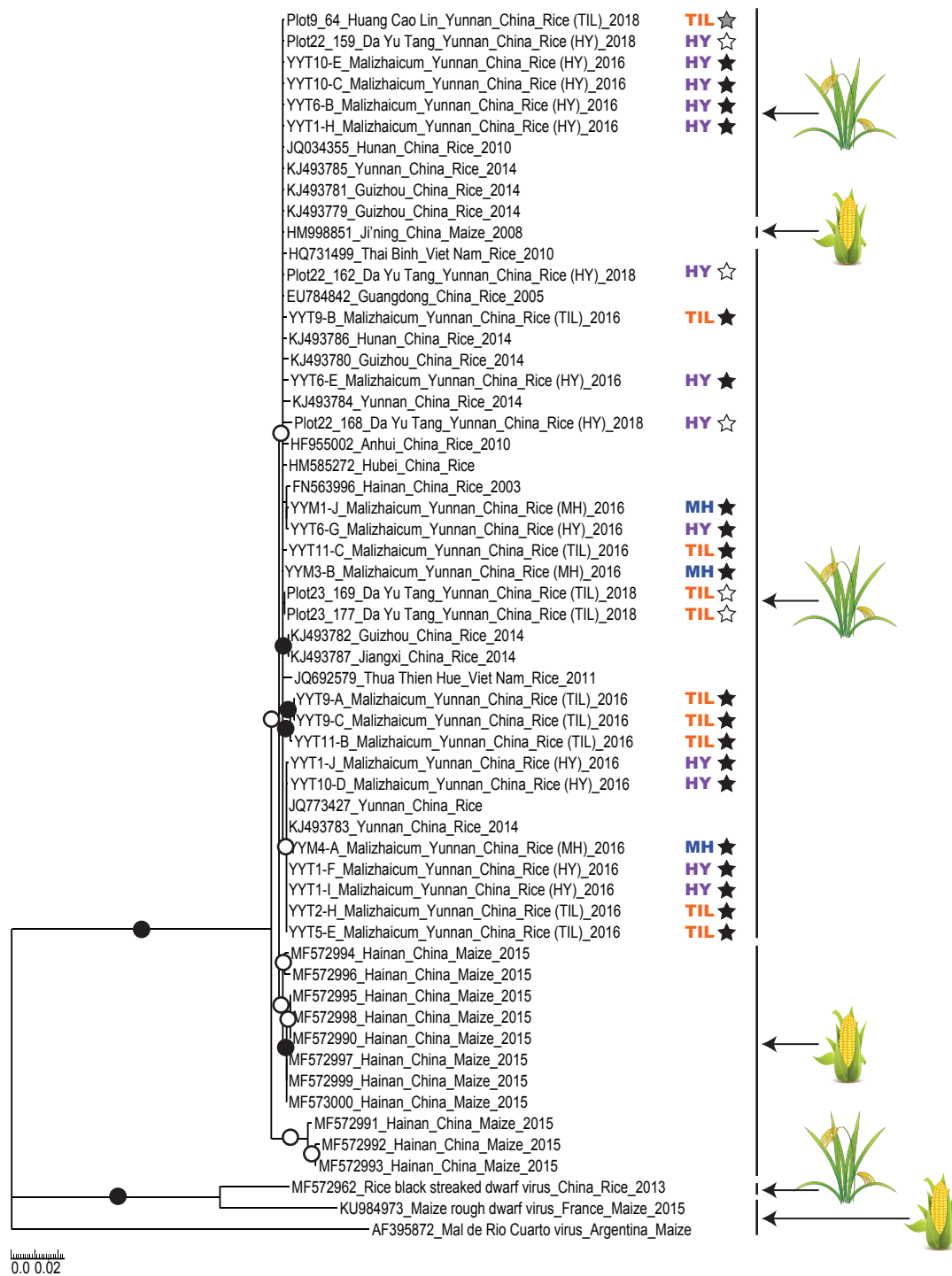


Figure 2. Maximum-likelihood phylogenetic trees of partial SRBSDV segment 8 sequences (624 nt in length). Branches with less than 50% bootstrap support were collapsed. Branches associated with a black dot have bootstrap supports above 90% whereas those with white dots have bootstrap supports above 70%. Variety group of the samples, the village where the plant sample was collected and the host plant are depicted on the right of the phylogenetic tree. Variety groups: traditional *indica* landraces group (TIL); Hongyang improved modern variety group (HY) and modern hybrid variety group (MH). Villages: grey star, Huangcaolin village; black star, Malizhai village and white star, Dayutang village.

3.4. The Emergence of SRBSDV and Long-Term HHRTS Sustainability

Contrary to past observations and predictions [10], we have revealed that a viral disease has emerged and become endemic in the HHRTS. Our study also revealed that the prevalence of the virus that causes this disease, SRBSDV, is significantly higher in the HY variety group than in the MH variety group. In addition, while not significant, the prevalence of SRBSDV in HY varieties is also higher than the prevalence in TIL varieties. This observation is consistent with a recent survey which found that farmers who since 2014 had shifted to using high-yielding HY varieties had perceived a subsequent clear increase in disease prevalence and severity [13]. Our results suggest that the rice black-streaked dwarf disease may have been responsible for this perception of “diseased plants” by the interviewed farmers.

Given that rice black-streaked dwarf disease can cause 30–50% rice yield losses [1,2], our study emphasizes both the need to better monitor the disease in the HHRTS, and the need to start considering ways to reduce its burden on rice production within the terraces. Our finding that different rice variety groups are differentially infected by SRBSDV is consistent with the notion that continuing with the traditional use of highly genetic diverse rice landraces in the HHRTS should reduce the overall impact of SRBSDV within the system. Specifically, the absence of detectable SRBSDV infections in several of the TIL, MH and HY varieties (Table 1) may indicate that the rice genetic resources for minimizing the impacts of SRBSDV on the HHRTS are already at hand. Continuing the HHRTS tradition of farmer-led selection of TIL and other varieties that are least affected by the virus may eventually protect the system from this virus.

A major challenge will be the maintenance of the secular HHRTS farm seed exchange organizations that rely on a variety of non-market mechanisms such as gifts, social reciprocity networks and shared infrastructure resources to enable the dissemination of TIL varieties [13]. In addition, our study shows that several MH and HY SRBSDV-tolerant rice varieties, could also be selected and used in the HHRTS. However, more studies are needed to assess the long-term susceptibility of these rice varieties, to monitor the changing impacts of rice black-streaked dwarf disease on yields in the HHRTS, and to better estimate the optimal proportions of modern vs. TIL varieties that should be deployed in the HHRTS to ensure both the maximization of rice yields and the continuing development of new TIL varieties within the terraces.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1999-4915/11/11/985/s1>, Figure S1: (A) Map of Asia and the location of the Honghe Hani rice terraces system (HHRTS). (B) Map of the Malizhai River Basin from HHRTS and location of the seven villages where rice plants were collected in 2016 and 2018. (C) Map of the Malizhai village and location of the rice fields were collected in 2016.

Author Contributions: X.H., J.-B.M., C.V. and P.R. conceived and designed the experiments; R.F., B.J., H.A., E.F. (Elisabeth Fournier), H.H., C.V. and P.R. collected the rice samples; P.A., O.M., P.-J.S., P.M., J.F., L.B., R.F., E.F. (Emmanuel Fernandez), C.J., A.D., C.V. and P.R. performed the experiments; P.A., O.M., P.M., D.F., J.F., V.G., D.P.M., P.G. and P.R. analyzed the data; P.A., J.-B.M., H.H., P.G., D.P.M. and P.R. wrote the paper that was edited and approved by all authors.

Funding: This study was funded by the Agropolis Fondation (E-Space flagship program) grant number 1504-004. D.P.M. has received a research grant from the National Research Foundation of South Africa.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lv, M.F.; Xie, L.; Wang, H.F.; Wang, H.D.; Chen, J.P.; Zhang, H.M. Biology of Southern rice black-streaked dwarf virus: A novel fijivirus emerging in East Asia. *Plant Pathol.* **2017**, *66*, 515–521. [CrossRef]
2. Zhou, G.; Xu, D.; Xu, D.; Zhang, M. Southern rice black-streaked dwarf virus: A white-backed planthopper-transmitted fijivirus threatening rice production in Asia. *Front. Microbiol.* **2013**, *4*, 270. [CrossRef] [PubMed]
3. Hoang, A.T.; Zhang, H.M.; Yang, J.; Chen, J.P.; Hebrard, E.; Zhou, G.H.; Vien, V.N.; Cheng, J.A. Identification, Characterization, and Distribution of Southern rice black-streaked dwarf virus in Vietnam. *Plant Dis.* **2011**, *95*, 1063–1069. [CrossRef] [PubMed]

4. Matsukura, K.; Towata, T.; Sakai, J.; Onuki, M.; Okuda, M.; Matsumura, M. Dynamics of Southern rice black-streaked dwarf virus in Rice and Implication for Virus Acquisition. *Phytopathology* **2013**, *103*, 509–512. [\[CrossRef\]](#)
5. Zhang, H.M.; Yang, J.; Chen, J.P.; Adams, M.J. A black-streaked dwarf disease on rice in China is caused by a novel fijivirus. *Arch. Virol.* **2008**, *153*, 1893–1898. [\[CrossRef\]](#)
6. Zhou, Y.C.; Nossoussou, M.; Kon, T.; Rojas, M.R.; Jiang, H.; Chen, L.F.; Gamby, K.; Foster, R.; Gilbertson, R.L. Evidence of local evolution of tomato-infecting begomovirus species in West Africa: Characterization of tomato leaf curl Mali virus and tomato yellow leaf crumple virus from Mali. *Arch. Virol.* **2008**, *153*, 693–706. [\[CrossRef\]](#)
7. Wang, Q.A.; Yang, J.A.; Zhou, G.H.; Zhang, H.M.; Chen, J.P.; Adams, M.J. The Complete Genome Sequence of Two Isolates of Southern rice black-streaked dwarf virus, a New Member of the Genus Fijivirus. *J. Phytopathol.* **2010**, *158*, 733–737. [\[CrossRef\]](#)
8. Zhou, Y.; Zhang, X.; Wang, D.; Weng, J.; Di, H.; Zhang, L.; Dong, L.; Zhang, H.; Zu, H.; Li, X.; et al. Differences in Molecular Characteristics of Segment 8 in Rice black-streaked dwarf virus and Southern rice black-streaked dwarf virus. *Plant Dis.* **2018**, *102*, 1115–1123. [\[CrossRef\]](#)
9. Bernardo, P.; Charles-Dominique, T.; Barakat, M.; Ortet, P.; Fernandez, E.; Filloux, D.; Hartnady, P.; Rebelo, T.A.; Cousins, S.R.; Mesleard, F.; et al. Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *ISME J.* **2018**, *12*, 173–184. [\[CrossRef\]](#)
10. Jiao, Y.M.; Li, X.Z.; Liang, L.H.; Takeuchi, K.; Okuro, T.; Zhang, D.D.; Sun, L.F. Indigenous ecological knowledge and natural resource management in the cultural landscape of China's Hani Terraces. *Ecol. Res.* **2012**, *27*, 247–263. [\[CrossRef\]](#)
11. Liao, J.; Huang, H.; Meusnier, I.; Adreit, H.; Ducasse, A.; Bonnot, F.; Pan, L.; He, X.; Kroj, T.; Fournier, E.; et al. Pathogen effectors and plant immunity determine specialization of the blast fungus to rice subspecies. *eLife* **2016**, *5*. [\[CrossRef\]](#)
12. Yang, L.; Liu, M.C.; Lun, F.; Yuan, Z.; Zhang, Y.X.; Min, Q.W. An Analysis on Crops Choice and Its Driving Factors in Agricultural Heritage Systems-A Case of Honghe Hani Rice Terraces System. *Sustainability* **2017**, *9*. [\[CrossRef\]](#)
13. Dedeurwaerdere, T.; Hannachi, M. Socio-economic drivers of coexistence of landraces and modern crop varieties in agro-biodiversity rich Yunnan rice fields. *Ecol. Econ.* **2019**, *159*, 177–188. [\[CrossRef\]](#)
14. Ben Saad, R.; Fabre, D.; Mieulet, D.; Meynard, D.; Dingkuhn, M.; Al-Doss, A.; Guiderdoni, E.; Hassairi, A. Expression of the *Aeluropus littoralis* ALSAP gene in rice confers broad tolerance to abiotic stresses through maintenance of photosynthesis. *Plant Cell Environ.* **2012**, *35*, 626–643. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Risterucci, A.M.; Hippolyte, I.; Perrier, X.; Xia, L.; Caig, V.; Evers, M.; Huttner, E.; Kilian, A.; Glaszmann, J.C. Development and assessment of Diversity Arrays Technology for high-throughput DNA analyses in *Musa*. *Tag. Theor. Appl. Genetics. Theor. Und Angew. Genet.* **2009**, *119*, 1093–1103. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **2011**, *6*, e19379. [\[CrossRef\]](#)
17. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [\[CrossRef\]](#)
18. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [\[CrossRef\]](#)
19. Huson, D.H.; Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **2006**, *23*, 254–267. [\[CrossRef\]](#)
20. Palanga, E.; Filloux, D.; Martin, D.P.; Fernandes, E.; Gargani, D.; Ferdinand, R.; Zabre, J.; Bouda, Z.; Neya, J.B.; Sawadogo, M.; et al. Metagenomic-based screening and molecular characterization of cowpea-infecting viruses in Burkina Faso. *PLoS ONE* **2016**, *11*, e0165188. [\[CrossRef\]](#)
21. Wang, Z.C.; Yu, L.; Jin, L.H.; Wang, W.L.; Zhao, Q.; Ran, L.L.; Li, X.Y.; Chen, Z.; Guo, R.; Wei, Y.T.; et al. Evaluation of Rice Resistance to Southern Rice Black-Streaked Dwarf Virus and Rice Ragged Stunt Virus through Combined Field Tests, Quantitative Real-Time PCR, and Proteome Analysis. *Viruses* **2017**, *9*. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Edgar, R.C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* **2004**, *5*, 113. [\[CrossRef\]](#) [\[PubMed\]](#)

23. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
24. Muhire, B.M.; Varsani, A.; Martin, D.P. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS ONE* **2014**, *9*, e108277. [[CrossRef](#)] [[PubMed](#)]
25. Guindon, S.; Delsuc, F.; Dufayard, J.F.; Gascuel, O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol. Biol.* **2009**, *537*, 113–137. [[CrossRef](#)]
26. Frichot, E.; Mathieu, F.; Trouillon, T.; Bouchard, G.; Francois, O. Fast and efficient estimation of individual ancestry coefficients. *Genetics* **2014**, *196*, 973–983. [[CrossRef](#)] [[PubMed](#)]
27. Carpi, G.; Holmes, E.C.; Kitchen, A. The evolutionary dynamics of bluetongue virus. *J. Mol. Evol.* **2010**, *70*, 583–592. [[CrossRef](#)]
28. Hon, C.C.; Lam, T.Y.; Drummond, A.; Rambaut, A.; Lee, Y.F.; Yip, C.W.; Zeng, F.; Lam, P.Y.; Ng, P.T.; Leung, F.C. Phylogenetic analysis reveals a correlation between the expansion of very virulent infectious bursal disease virus and reassortment of its genome segment B. *J. Virol.* **2006**, *80*, 8503–8509. [[CrossRef](#)]
29. Lu, Z.; Liu, H.; Fu, S.; Lu, X.; Dong, Q.; Zhang, S.; Tong, S.; Li, M.; Li, W.; Tang, Q.; et al. Liao ning virus in China. *Virol. J.* **2011**, *8*, 282. [[CrossRef](#)]
30. Matthijssens, J.; Heylen, E.; Zeller, M.; Rahman, M.; Lemey, P.; Van Ranst, M. Phylodynamic analyses of rotavirus genotypes G9 and G12 underscore their potential for swift global spread. *Mol. Biol. Evol.* **2010**, *27*, 2431–2436. [[CrossRef](#)]
31. Stenger, D.C.; Sisterson, M.S.; French, R. Population genetics of Homalodisca vitripennis reovirus validates timing and limited introduction to California of its invasive insect host, the glassy-winged sharpshooter. *Virology* **2010**, *407*, 53–59. [[CrossRef](#)] [[PubMed](#)]
32. Chao, L.; Rang, C.U.; Wong, L.E. Distribution of spontaneous mutants and inferences about the replication mode of the RNA bacteriophage phi6. *J. Virol.* **2002**, *76*, 3276–3281. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

2.4. Section 3: Heterogeneity of the rice microbial community of the Chinese centuries-old Honghe Hani rice terraces system

Heterogeneity of the rice microbial community of the Chinese centuries-old Honghe Hani rice terraces system

Pascal Alonso,^{1,2} Laurence Blondin,^{1,2}
Pierre Gladieux,^{2,3} Frédéric Mahé,^{1,2} Hervé Sanguin,^{1,2}
Romain Ferdinand,^{1,2} Denis Filloux,^{1,2} Eric Desmarais,⁴
Frédérique Cerqueira,⁴ Baihui Jin,⁵ Huichuan Huang,⁵
Xiahong He,^{5,6} Jean-Benoit Morel,^{2,3} Darren P. Martin,⁷
Philippe Roumagnac^{1,2} and Christian Vernière^{1,2*}

¹CIRAD, BGPI, Montpellier, France.

²BGPI, INRAE, CIRAD, Institut Agro, Univ Montpellier, Montpellier, France.

³INRA, BGPI, Montpellier, France.

⁴ISEM, CNRS, University of Montpellier, IRD, EPHE, Montpellier, France.

⁵State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan Agricultural University, Kunming, 650201, China.

⁶Southwest Forestry University, Kunming, China.

⁷Computational Biology Group, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, 4579, South Africa.

Summary

The Honghe Hani rice terraces system (HHRTS) is a traditional rice cultivation system where Hani people cultivate remarkably diverse rice varieties. Recent introductions of modern rice varieties to the HHRTS have significantly increased the severity of rice diseases within the terraces. Here, we determine the impacts of these recent introductions on the composition of the rice-associated microbial communities. We confirm that the HHRTS contains a range of both traditional HHRTS landraces and introduced modern rice varieties and find differences between the microbial communities of these two groups. However, this introduction of modern rice varieties has not strongly impacted the overall diversity of the HHRTS rice microbial community. Furthermore, we find that the rice varieties (i.e. groups of closely related genotypes)

have significantly structured the rice microbial community composition (accounting for 15%–22% of the variance) and that the core microbial community of HHRTS rice plants represents less than 3.3% of all the microbial taxa identified. Collectively, our study suggests a highly diverse HHRTS rice holobiont (host with its associated microbes) where the diversity of rice hosts mirrors the diversity of their microbial communities. Further studies will be needed to better determine how such changes might impact the sustainability of the HHRTS.

Introduction

During and since the Green Revolution, governments and agricultural stakeholders all around the world have promoted the widespread introduction to cropping systems of modern high-yielding crop varieties. By contributing to the simplification and homogenization of cropping systems this trend has likely also inadvertently contributed to increases in the occurrence of various crop diseases (Keesing *et al.*, 2006; Stuckenberg and McDonald, 2008; Roossinck and Garcia-Arenal, 2015; Bernardo *et al.*, 2018). Nevertheless, there are a number of Globally Important Agricultural Heritage Systems (GIAHS) within which landraces (traditional varieties) are still cultivated (FAO, <http://www.fao.org/giahs/en/>). Although many of these GIAHSs were left largely unchanged by the Green Revolution, some, such as the Honghe Hani rice terraces system (HHRTS), have more recently experienced both the introduction of modern high-yielding rice varieties and an increase in chemical usage (Yang *et al.*, 2017; Dedeurwaerdere and Hannachi, 2019).

The HHRTS, which was recently listed as a World Cultural Heritage Site (UNESCO, 2013), is a renowned rice terrace landscape where remarkably diverse rice varieties have been cultivated for over 1300 years by the Hani people (Jiao *et al.*, 2012). The HHRTS consists of a collection of unique man-made vertically structured ecological landscapes comprising cascading terraced rice fields sandwiched between forests and villages at the top, and a river at the bottom (Cui *et al.*, 2008; Yang *et al.*, 2017). Hundreds of years of rice varietal selection within the HHRTS

Received 3 March, 2020; revised 29 May, 2020; accepted 2 June, 2020. *For correspondence. E-mail christian.verniere@cirad.fr; Tel. +33 (0) 4 99 62 48 58; Fax +33 (0)4 99 62 48 48.

has yielded at least 195 local rice landraces (*Oryza sativa*) and 47 wild rice landraces (*Oryza rufipogon* and *Oryza nivara*) (Jiao *et al.*, 2012). These diverse rice landraces are grown by each Hani household in complex heterogeneous mosaics, in narrow paddy fields averaging $\sim 150 \text{ m}^2$ within each of which only one rice landrace is cultivated.

While increased diversification of rice varieties has likely reduced the severity of rice-associated diseases (Zhu *et al.*, 2000), it is usually hypothesized that the structure of the HHRTS has likely also limited the occurrence and spread of diseases within the system. For example, the spread of *Pyricularia oryzae*, the causal agent of rice blast, is apparently hampered by the high diversity of basal and effector-triggered immune responses that are displayed by the *japonica* and *indica* rice varieties commonly cultivated within the HHRTS (Liao *et al.*, 2016).

However, a recent socio-economic survey has revealed a significant increase in the severity of diseases occurring within HHRTS fields where the government-promoted 'HongYang' improved rice variety has been cultivated (Dedeurwaerdere and Hannachi, 2019). Increased use of this improved variety since 2010 may, for example, account for increases over the past decade in the prevalence within the HHRTS of Southern rice black-streaked dwarf virus (SRBSDV) (Alonso *et al.*, 2019).

Besides the role played in plant health by intrinsic plant mechanisms, plant-associated microbial communities are also likely to be directly or indirectly involved in plant health and plant development (Berendsen *et al.*, 2012; Hacquard *et al.*, 2017; Hassani *et al.*, 2019; Vannier *et al.*, 2019). It is now widely accepted that the structuring of plant microbiota is controlled by plant-specific factors (Vorholt, 2012; Bulgarelli *et al.*, 2013; Reinhold-Hurek *et al.*, 2015; Hamonts *et al.*, 2018) and that plant genotype may further impact microbial communities (Bulgarelli *et al.*, 2015; Sapkota *et al.*, 2015; Wagner *et al.*, 2016). Specifically, plant domestication or selection has possibly contributed to a small but significant change in plant microbial communities (Redford *et al.*, 2010; Peiffer and Ley, 2013; Bouffaud *et al.*, 2014; Ofek-Lalzar *et al.*, 2014; Bulgarelli *et al.*, 2015; Edwards *et al.*, 2015). The HHRTS is therefore a good potential candidate for determining the impacts of cultivating modern rice varieties on aspects of traditional rice cultivation systems: impacts such as the composition of their associated microbial communities. Overall, modern varieties have low genetic diversity due to the recurrent selection for traits contributing to, among other things, plant yield, rice quality, and resistance to biotic and abiotic stresses (Meyer and Purugganan, 2013). This genetic simplification has often been implicated in reducing the bacterial and fungal diversity in the rhizosphere and phyllosphere of modern cultivated crop varieties relative to that found in the rhizosphere and phyllosphere of wild species and landraces (reviewed in Cordovez *et al.*, 2019).

For example, modern cultivated pea and broad bean varieties have less promiscuous interactions with symbionts than do their wild relatives (Mutch and Young, 2004). Similarly, the bacterial and fungal diversity in the rhizosphere and phyllosphere of uncultivated agave species is higher than that of cultivated agave (Coleman-Derr *et al.*, 2016). Small but significant domestication effects have also been noted when comparing the root microbial communities of wild and modern varieties of barley (Bulgarelli *et al.*, 2015; Szoboszlay *et al.*, 2015). Two potential exceptions to this trend are modern lettuce and sunflower varieties, which harbour degrees of rhizobacterial diversity that are similar to, or in some cases higher than, those of their uncultivated relatives (Cardinale *et al.*, 2015; Leff *et al.*, 2017).

The microbial community associated with the rice rhizosphere has been the most intensively investigated (Edwards *et al.*, 2015; Wang *et al.*, 2016; Edwards *et al.*, 2018; Moronta-Barrios *et al.*, 2018; Ding *et al.*, 2019). Specifically, several studies have focused on the comparison of the rhizosphere or seed-associated microbial communities of wild *Oryza*, rice landraces and modern rice varieties (Shenton *et al.*, 2016; Shi *et al.*, 2019; Kim *et al.*, 2020). For instance, Shenton and colleagues (2016) have shown that the rhizosphere bacterial community of wild *Oryza* differed substantially from that of cultivated rice in terms of both species richness and composition (Shenton *et al.*, 2016). Interestingly, this study also revealed that landraces had a rhizosphere bacterial community composition that was intermediate between that of wild *Oryza* and cultivated rice varieties (Shenton *et al.*, 2016). It is also apparent that fungal communities differ more substantially than bacterial communities between the seed or rhizosphere microbial communities of cultivated varieties and wild *Oryza* (Shi *et al.*, 2019; Kim *et al.*, 2020).

Here, we hypothesized that the introduction of modern rice varieties to the HHRTS may have impacted rice plant microbial communities within the HHRTS. We initially determined the degree of rice genetic diversity within the paddy fields of a single HHRTS village, Malizhai that has adopted a mixed landrace/modern variety system. Then, we compared the diversity and composition of the microbial communities of introduced modern rice varieties and HHRTS landraces.

Results and discussion

Both landraces and modern rice varieties are grown in the Malizhai HHRTS

Malizhai is a village broadly representative of the HHRTS throughout the Hani region (Jiao *et al.*, 2012). To obtain a detailed understanding of rice genetic diversity within the Malizhai HHRTS, we sampled 19 rice paddies, each in a 2 km^2 area of Malizhai. While 11 of the rice paddies from

which samples were collected were referred to as 'traditional varieties' by the Malizhai farmers, eight were referred to as 'modern varieties'. We performed a genotyping by sequencing (GBS) analysis of the 19 sampled rice paddies (hereafter referred to as our Malizhai GBS dataset) and compared these to a reference dataset containing whole-genome re-sequencing data produced by the 3000 Rice Genomes Project (Wang *et al.*, 2018) from nine rice accessions randomly selected from among four representative subpopulations of the *indica* rice subspecies (subpopulations XI-1A, XI-1B, XI-2 and XI-3).

This GBS comparison revealed that only one of the groups of Malizhai rice genotypes belonged to the *japonica* rice subspecies with the remaining 18 group of rice genotypes belonging to the *indica* rice subspecies (Fig. 1A, B). Whereas 10 of the 18 group of Malizhai *indica* genotypes were identified by farmers as being 'traditional varieties' and will hereafter be referred to as the 'group of HHRTS landraces', the other eight genotypes were identified by farmers as being 'modern varieties' and will hereafter be referred to as the 'group of modern rice'. Our inference of population subdivision by partitioning genotypes into K ancestral populations, based both on the cross-entropy criterion (Fig. S1) and visual analysis of clustering patterns (Fig. 1B), revealed that the model with $K = 4$ clusters captured most of the structure in the data. Hence, at $K = 4$, besides the *indica* rice genotypes from the 3000 Rice Genomes Project (coloured in blue, Fig. 1B), three groups from the Malizhai HHRTS were distinguishable, including a group of *japonica* genotypes (coloured in orange, Fig. 1B), the group of HHRTS landraces (coloured in green, Fig. 1B), and the group of modern varieties (coloured in brown, Fig. 1B). Our GBS analysis confirmed that the rice genotypes are heterogeneously distributed in the Malizhai HHRTS forming a mosaic landscape of modern genotypes and landraces (Fig. 1C). The eight groups of genotypes representing the group of modern varieties included commercial modern hybrid F1 varieties that are widely and intensively cultivated in the rice-growing plains of China (Table 1). Nucleotide diversity was slightly lower within the group of modern varieties ($\pi = 8.05 \times 10^{-4}$) than that of the HHRTS landraces group ($\pi = 9.43 \times 10^{-4}$).

The model with $K = 8$ clusters also revealed patterns of clustering that seem biologically sensible. At $K = 8$, the group of modern varieties could be further split into two subgroups, referred to as the modern rice subgroup1 (samples from fields YYM1, YYM3, YYM6, YYM8 and YYM9, Fig. 1B) and the modern rice subgroup2 (samples from fields YYM4, YYM5 and YYM7, Fig. 1B). Unexpectedly, but in accordance with farmer assignments, the government-promoted HongYang varieties grown in three fields (YYT1, YYT6 and YYT10; Table 1) that were believed to be 'modern improved varieties' clearly share a recent ancestry with the HHRTS landraces (Fig. 1A). This implies that HongYang is based

on a traditional landrace that likely originated from the group of HHRTS landraces (Dedeurwaerdere and Hannachi, 2019). Clustering patterns at $K = 8$ also revealed that the group of HHRTS *indica* landraces can be further subdivided into four subgroups, which are referred to as HHRTS landraces subgroup1 (YYT1, YYT6 and YYT10, i.e. HongYang varieties), HHRTS landraces subgroup2 (YYT2, YYT11), HHRTS landraces subgroup3 (YYT3, YYT5) and HHRTS landraces subgroup4 (YYT7, YYT9) (Fig. 1B).

We have thus found that rice varieties grown in the Malizhai HHRTS fall into two main genetic groups; one including newly introduced commercial Hybrid F1 modern varieties and the other including several traditional or improved (HongYang) varieties. This confirms that changes in cultural practices in the Malizhai HHRTS, and likely in other villages as well, have resulted in a shift in the genetic makeup of cultivated rice in the terraces. We also find evidence that further replacement of HHRTS landraces by modern varieties could slightly reduce the over-all genetic diversity of rice within the HHRTS.

The microbial communities of rice roots and stems are highly divergent

Having characterized the genetic diversity of rice grown in the Malizhai HHRTS, we next sought to characterize the bacterial and fungal components of the microbial communities of sampled rice plants. Our samples included the microbes living at the surface and within plant tissue. Based on the analysis of GBS and whole-genome re-sequencing data, we split the sampled *indica* rice plants into a group of modern varieties (including 78 plants from eight different modern varieties) and a group of HHRTS landraces (including 98 plants from 10 different traditional or improved HHRTS varieties). For consistency, we excluded plants of the YYT8 *japonica* rice subspecies (Table 1) from this analysis. We then examined the stem and root microbial communities of plants belonging to the HHRTS landrace and modern variety groups.

Metabarcoding of bacterial communities produced after the bioinformatics processing 120 666 high-quality reads from the root compartment and 92 400 from the stem compartment respectively corresponding to 325 and 91 Operational taxonomic units (OTUs) with each OTU representing over 50% prevalence per rice variety (Table S1). For the fungal communities, we obtained 30 858 high-quality reads for the root compartment and 21 750 high-quality reads for the stem for the root compartment respectively representing 110 and 105 OTUs with each OTU representing over 50% prevalence per rice variety (Table S1). Details of the number of sequences and OTUs recovered after each step of the bioinformatics processing are available in Table S1. Given that too few samples from the field YYT4 yielded enough high-quality reads to reach the rarefaction

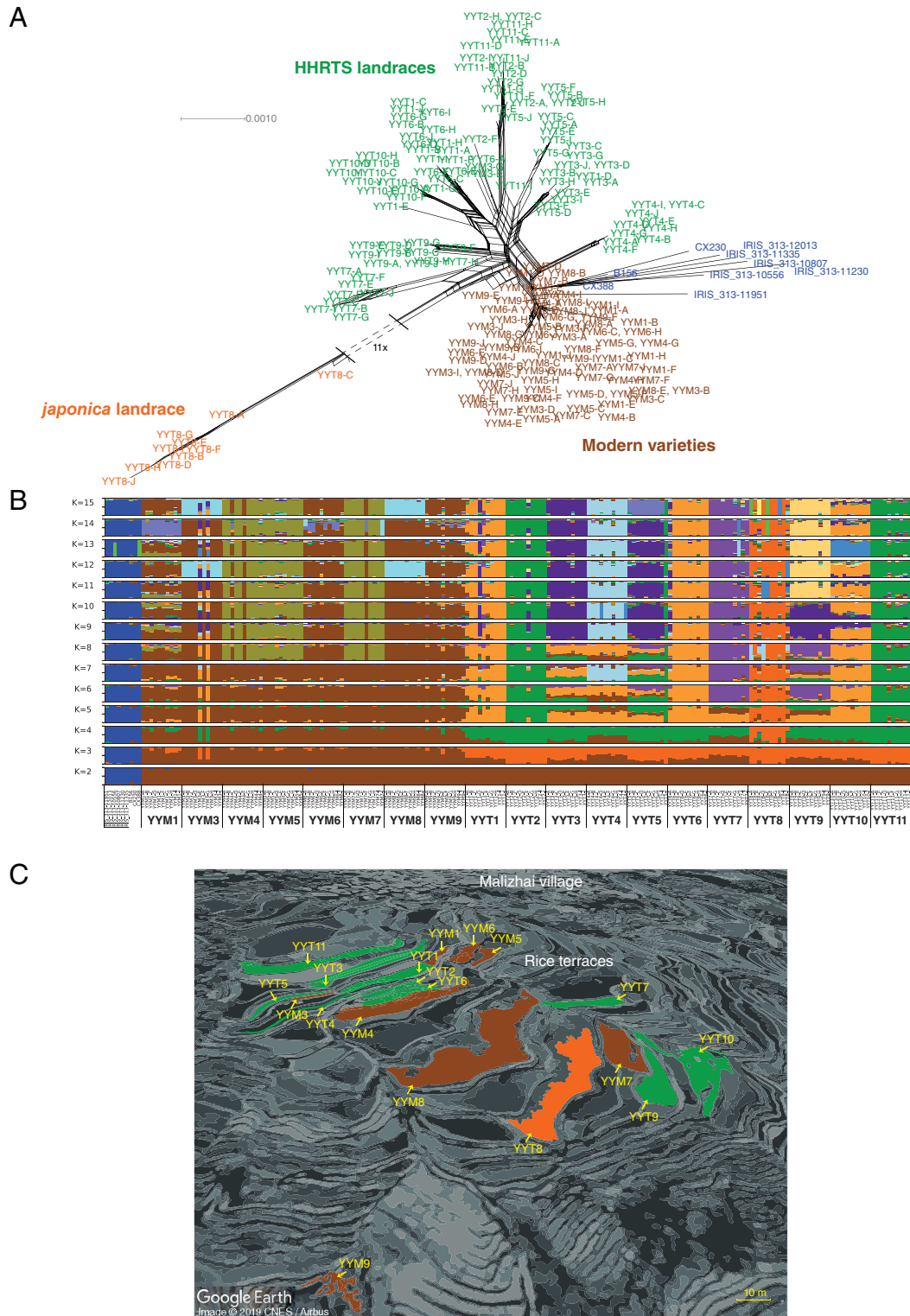


Fig 1. A. Neighbour-Net split decomposition network indicating the relationships between rice accessions based on 10 028 analysed SNPs. Plant samples assigned to the ‘HHRTS landraces’, the ‘modern rice varieties’, the *japonica* subspecies and representative subset of *indica* subgroups from the 3000 Rice Genomes Project are labelled in green, brown, orange and blue respectively. B. Ancestry proportions inferred using sNMF for models with $K = 2$ to $K = 15$ ancestral populations. Each accession is represented by a vertical bar divided into K segments of different colours, representing proportions of ancestry in K ancestral populations for a single accession, with colours corresponding to ancestral populations. C. Map of the Malizhai rice terraces and location of the rice fields that were collected in 2016. Terraces were HHRTS landraces, modern rice varieties and *japonica* varieties, which are coloured in green, brown and orange respectively.

Table 1. Genotyping of rice varieties from the Yuanyang rice terraces of China and assignment of rice varieties to three rice genetic groups ('HHRTS landraces group', 'Modern rice group' and *Oryza sativa* subspecies *japonica*).

Sampling field	Rice genetic group	Rice variety name(given by the farmers)	Number of samples
YYT1	HHRTS landraces group (HongYang)	Chepugu	9
YYT2	HHRTS landraces group	Lubaigu	10
YYT3	HHRTS landraces group	Zaogu	10
YYT4	HHRTS landraces group	Nuogu	10
YYT5	HHRTS landraces group	Epugu	10
YYT6	HHRTS landraces group (HongYang)	Chepugu	9
YYT7	HHRTS landraces group	Jianshuigu	10
YYT8	<i>Oryza sativa</i> ssp. <i>japonica</i>	Honglueduolu	10
YYT9	HHRTS landraces group	Nuogu	10
YYT10	HHRTS landraces group (HongYang)	Jinpinggu	10
YYT11	HHRTS landraces group	Luhonggu	10
YYM1	Modern rice group	Mingliangyou 527	10
YYM3	Modern rice group	Mingliangyou 528	8
YYM4	Modern rice group	Hefeng 177	10
YYM5	Modern rice group	Zhongyou 177	10
YYM6	Modern rice group	Liangyou 2186	10
YYM7	Modern rice group	Guofeng 1	10
YYM8	Modern rice group	Liangyou 725	10
YYM9	Modern rice group	Liangyou 2161	10

thresholds, this variety was further excluded from the group of HHRTS landraces in all subsequent analyses. Data for the group of HHRTS landraces were drawn from 88 plant samples belonging to nine different rice traditional varieties. The OTUs tables are available in Tables S2 and S3.

Root bacterial communities were significantly richer than those of the stem (Fig. S2A) and the structure and composition of these bacterial communities were also significantly different (Fig. S2C). Additionally, while the α diversity indices (measured by the richness index and Shannon index) of fungal root and stem communities were not significantly different (Fig. S2B), the fungal community compositions were significantly different between the stem and root compartments, albeit to a lower degree than that observed for the bacterial communities (Fig. S2C/D). These results remain consistent with previous observations in other species that the strongest determinant of the microbial community compositions is the compartment from which these are drawn (Knief *et al.*, 2012; Lundberg *et al.*, 2012; Coleman-Derr *et al.*, 2016).

Wide inter-individual variation in α diversity of microbial communities associated with HHRTS rice plants

The root and stem microbial community of HHRTS rice plants were subsequently studied separately to detect potential impacts on these of the shifting composition of rice plant populations that is currently occurring in the HHRTS (landraces vs. modern varieties). Alpha diversity indices were not significantly different between modern rice varieties and HHRTS landraces (Fig. 2) when considering either bacterial or fungal communities in either the stem or root

samples. Additionally, the six-subgroups partitioning revealed a low variability in alpha diversity indices among HHRTS rice subgroups even if the indices of α diversity between landraces subgroups 3 and 4 were significantly different (Fig. S3). It should be stressed, however, that the α diversity indices estimated across landraces and modern varieties, or for each of the six subgroups of landraces and modern varieties were highly variable (Fig. 2 and Fig. S3), suggesting that factors other than rice genotype may be shaping the richness of microbial communities at the individual rice plant level. Such wide inter-individual variations in α diversity are reminiscent of the inter-individual variability of microbial communities observed previously in human saliva. This variability in saliva has been interpreted as a possible consequence of the human mouth being exposed to highly variable external (food) and internal (digestion exchanges) environments, and being subjected to a wide variety of oral hygiene regimes (Hall *et al.*, 2017; Verma *et al.*, 2018). Accordingly, our results suggest that individual rice plants within a field may have encountered during their lifetimes a variety of physical and chemical environmental conditions that have left an imprint on the composition of their associated microbial communities. The impacts of heterogeneous micro-environments on rice-associated microbial community structures may be particularly pertinent to flooded rice systems in general (Fernández-Valiente and Quesada, 2004; Cui *et al.*, 2008; Xie *et al.*, 2011; Jiao *et al.*, 2012), and specifically to the HHRTS flooded rice system, where micro-environment variability could conceivably arise for a multitude of different reasons including (i) the high densities at which rice is planted; (ii) the compactness of the terraces; (iii) the low depth of the floodwaters within the terraces; (iv) the variability of ancillary food

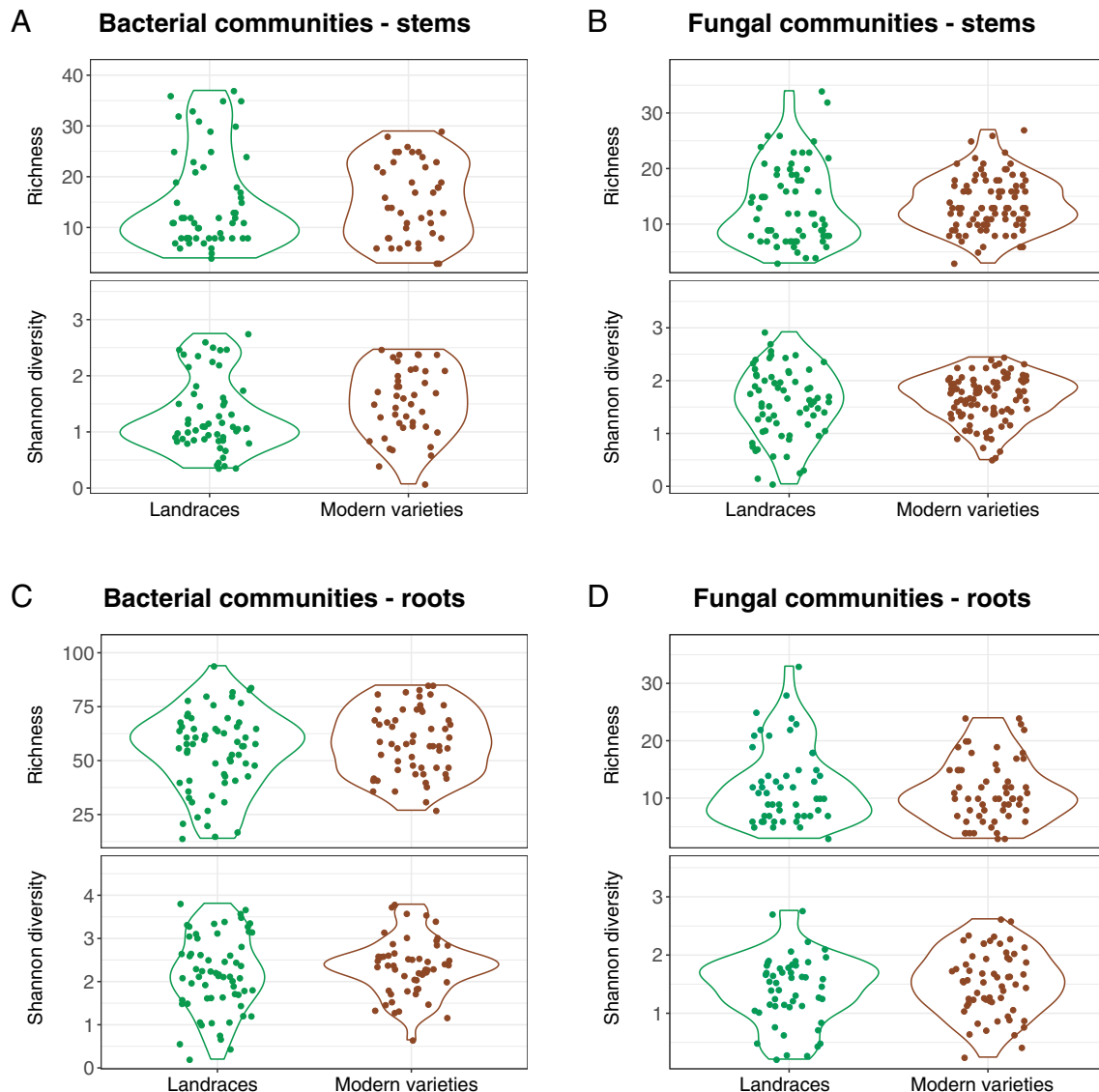


Fig 2. Violin plots of rice microbial α diversity (richness and Shannon diversity indexes) across HHRTS landraces and modern rice varieties for (A) the rice stem bacterial communities and (B) the rice stem fungal communities as well as (C) the rice root bacterial communities and (D) the rice root fungal communities. P -values of Tukey HSD tests were all >0.05 between landrace and modern communities whatever the community (bacteria or fungi) and the diversity indexes.

production systems that are co-located in the terraces which can include ducks, fish, frogs, and snails; and (v) erratic increases in the levels of suspended soil particles in the floodwaters arising as a consequence of periodic disturbances of silt by labourers and water buffalos within the paddies.

Microbial community structure of modern rice varieties and HHRTS landraces are slightly different

We further examined differences between the microbial community compositions associated with modern rice varieties and those associated with HHRTS landraces.

We first calculated the β -diversity and used a permutational multivariate analysis of variance (PERMANOVA) of microbial communities between modern rice varieties and HHRTS landraces for both stem and root samples using phylogeny-based UniFrac distances either weighted or unweighted by the abundance of OTUs. These PERMANOVA analyses revealed a slight (R^2 values ranged from 0.03 to 0.06), but significant effect of variety type for both stem and root bacterial and fungal communities using the unweighted UniFrac distances (Fig. 3). This finding supports the hypothesis that the introduction of modern rice varieties in the HHRTS could have caused a slight modification in the structure of rice-associated

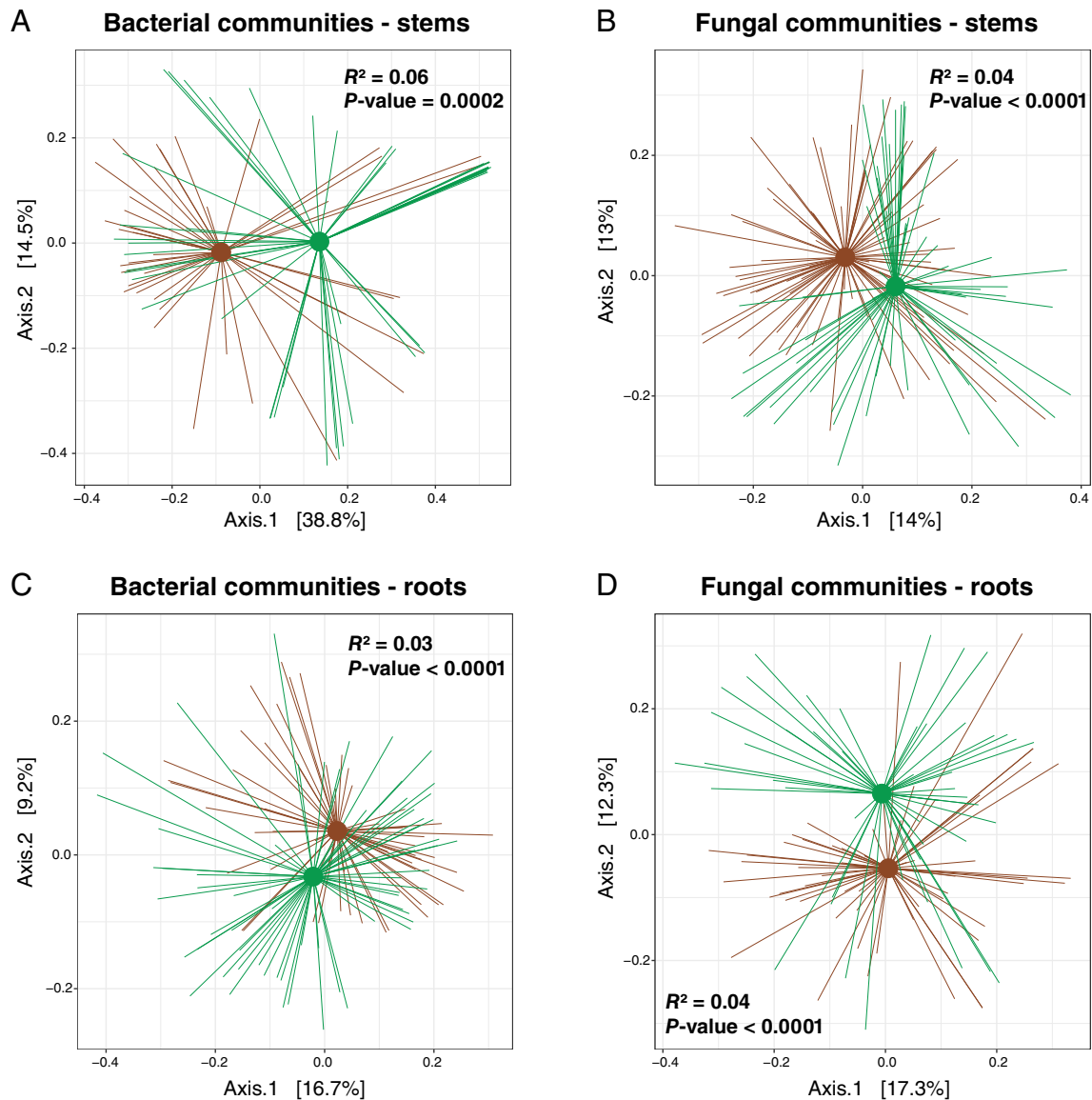


Fig 3. Principal coordinates analysis (PCoA) plots based on unweighted UniFrac distances of (A) rice stem bacterial communities and (B) rice stem fungal communities as well as (C) rice root bacterial communities and (D) rice root fungal communities. Communities from the HHRTS landraces and from the modern rice varieties are labelled in green and brown respectively. Axes represent the two dimensions explaining the greatest proportion of variance in the communities for each analysis. Permutational multivariate analysis of variance (PERMANOVA) results are indicated (R^2 and the P -value).

microbial communities. This slight effect is consistent with findings from recent studies on other plant species (Emmett *et al.*, 2017; Hamonts *et al.*, 2018; Compant *et al.*, 2019). In contrast, the quantitative weighted UniFrac distances did not support a difference between the structures of HHRTS landrace microbial communities and the modern variety microbial communities (Fig. S4). This result suggests that the slight microbial variations observed between both rice groups in the Malizhai HHRTS are primarily driven by differences in the distributions of rare taxa between the microbial community associated with the HHRTS landrace and modern variety groups.

Rice variety is a key factor of the structure of the microbial community in Malizhai HHRTS

We examined differences between the microbial community compositions associated with the six rice subgroups of landraces and modern varieties that were sampled in 2016 in the Malizhai HHRTS and identified using clustering algorithms assuming $K = 8$ clusters (Fig. 4). The PERMANOVA analysis revealed a significant effect of the rice subgroup (R^2 values ranged from 0.14 to 0.22, $P < 0.0001$) for both stem and root bacterial and fungal communities using the unweighted UniFrac distances (Fig. 4). This result suggests

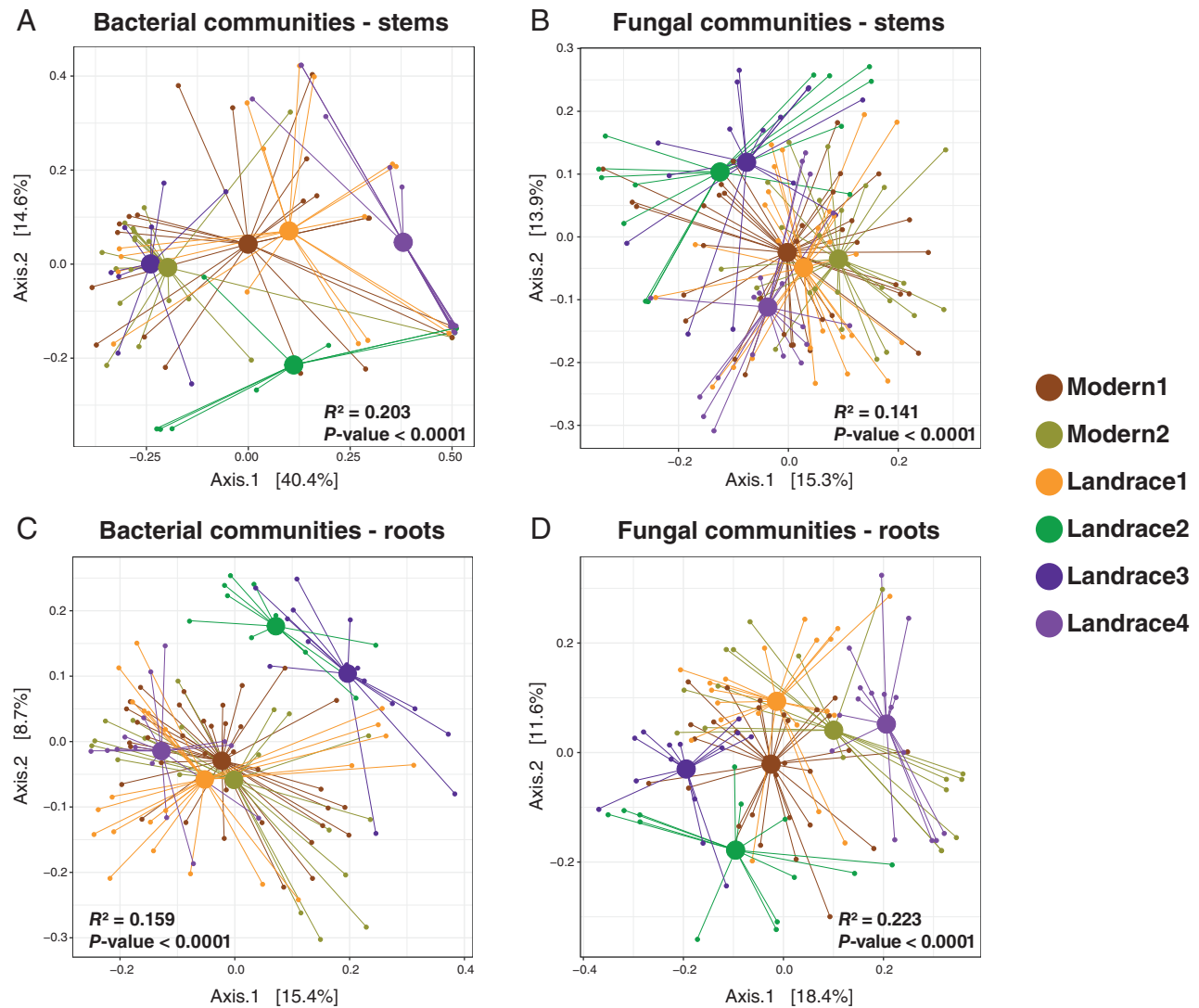


Fig 4. Principal coordinate analysis (PCoA) plots based on unweighted UniFrac distances of (A) rice stem bacterial communities, (B) rice stem fungal communities, (C) rice root bacterial communities and (D) rice root fungal communities. Communities associated with the six subgroups of HHRTS rice varieties are labelled in orange (HHRTS landraces subgroup1), green (HHRTS landraces subgroup2), blue (HHRTS landraces subgroup3), purple (HHRTS landraces subgroup4), brown (modern subgroup1) and olive green (modern subgroup2) respectively. Axes represent the two dimensions explaining the greatest proportion of variance in the communities for each analysis. Permutational multivariate analysis of variance (PERMANOVA) results are indicated (R^2 and the P -values).

that rice subgroups, each composed of related rice varieties, play a role in structuring the HHRTS microbial communities. This finding is in line with results from another study that revealed that genotypic differences in rice had a significant effect on root-associated microbial communities (Edwards *et al.*, 2015). Principal Coordinate Analyses (PCoA) based on unweighted UniFrac distances showed that bacterial communities from both stems and roots, and fungal communities from the stems of the modern subgroups and landrace subgroup1 (corresponding to the governmental-promoted HongYang varieties) were more similar to each other and farther from the three other landraces subgroups (Fig. 4). The pairwise PERMANOVA

analysis using the unweighted UniFrac distances further confirmed that the bacterial communities from roots were not significantly different between both modern rice subgroups and between the modern rice subgroup2 and the landrace subgroup1 (Table S4). This result suggests that varietal improvement for higher yields or pest resistance that was derived either from improved local landraces (HongYang varieties, landrace subgroup1) or from exogenous rice varieties (modern high-yielding rice hybrids) has led to a partial homogenization of microbial communities among the improved varieties grown in the Malizhai HHRTS. This convergence in microbial communities can be explained by shared selected features of improved rice

varieties such as their root morphology or their root exudation (or other traits) that have induced the assembly of similar microbial communities in the Malizhai HHRTS context (Szoboszlay *et al.*, 2015; Perez-Jaramillo *et al.*, 2018; Cordovez *et al.*, 2019). Based on these results, we tested whether partitioning the Malizhai HHRTS varieties into 'improved varieties' (including modern and HongYang varieties) and 'traditional landraces' (including landraces subgroups 2, 3 and 4) would have revealed a higher degree of population structuring in the Malizhai HHRTS microbial communities. The PERMANOVA analysis revealed a slight (R^2 values ranged from 0.04 to 0.05), but significant effect of variety type for both stem and root bacterial and fungal communities using the unweighted UniFrac distances, suggesting that the 'varietal improvement' factor explain less the structure of microbial communities than the 'rice genotype' factor.

We further compared the relative abundance of individual OTUs between the six rice subgroups of landraces and modern varieties using Wilcoxon signed-rank tests. We then used the Metacoder approach, which relies on a taxonomic tree-based visualization (called a heat tree), to illustrate whether the relative abundance of microbial OTUs differed significantly between the six rice subgroups (Figs S5, S6, S7 and S8). These analyses highlighted 243 OTUs with significantly different (with a P -value threshold <0.05) degrees of relative abundance between the six rice subgroups (Tables S5 and S6). Only two fungal OTUs of the 243 OTUs were more abundant in stems of all four rice subgroups of traditional and improved HHRTS varieties than in both modern rice subgroups (Table S6), suggesting a reduced core microbial community of rice varieties cultivated in the Malizhai HHRTS.

Phylogenetic distances between the microbial community and rice genotypes were correlated

Given that we found differences between the microbial communities of modern rice varieties and HHRTS landraces, we investigated whether more closely related genotypes tend to have more similar microbial communities. We conducted Mantel tests using the unweighted UniFrac distances of OTUs and genetic distances between rice genotypes. All four of the unweighted UniFrac distance matrices of microbial communities were significantly correlated with the genetic distances between the rice genotypes they are associated with (Table S7). This result is consistent with previous studies focusing on wild and modern rice varieties (Shenton *et al.*, 2016; Kim *et al.*, 2020), maize and other Poaceae (Bouffaud *et al.*, 2014) suggesting that the evolutionary history of poaceous crop plants has had an impact on the evolutionary history of their microbial communities.

The core microbial community associated with HHRTS rice plants encompassed a small number of taxa

We hypothesized that the 'core' microbial community in the stem and root of Malizhai HHRTS rice plants is limited in size. To test this hypothesis, we considered OTUs only present in at least 80% of the stem and root samples from plants belonging to the 18 *indica* varieties (including all the modern varieties and HHRTS landraces), with no consideration of the relative abundance of the taxa. The core stem microbial community consisted of just two fungal OTUs and one bacterial OTU and the core root microbial community consisted of just two fungal OTUs and three bacterial OTUs (Table 2). We further considered separately the stem and root core microbial communities of the modern rice varieties and HHRTS landraces. The core stem microbial community of modern rice varieties consisted of two fungal OTUs and the HHRTS landraces core stem microbial community included these two-same fungal OTUs and one additional bacterial OTU (Table 2 and Table S8). Whereas the core root microbial community of the modern rice varieties consisted of six bacterial OTUs and one fungal OTU that of the HHRTS landraces consisted of two bacterial OTUs and one fungal OTU (Table 2 and Table S8). Consequently, the core microbial community defined by our criteria ranged from 0.7% to 3.3% of the total number of taxa identified within and among the HHRTS landraces and modern rice varieties (Table 2). These low percentages are consistent with another study that identified a core microbial community constituted of 1.5% of the total OTUs associated with rice plants sampled from three Californian rice fields (Edwards *et al.*, 2015). The relative abundances of the 'core' OTUs associated with the Malizhai HHRTS rice varieties accounted for between 0.7% and 26.4% of the total number of assigned sequencing reads (see relative abundances in Table S8). The core microbial community, therefore, consisted of a combination of rare and abundant OTUs (Table S8). Besides revealing a high heterogeneity of HHRTS microbial community, our study also indicates that the HHRTS rice holobiont [host with its associated microbes that can potentially affect the phenotypes of rice plants (Vandenkoornhuyse *et al.*, 2015; Theis *et al.*, 2016)] is highly diverse with microbial communities mirroring the diversity their rice hosts.

Towards disentangling the biotic and abiotic factors affecting the Malizhai HHRTS rice microbial community

Our results show that the population structure of Malizhai HHRTS rice varieties is one of the factors that impact the community structure of associated microbes. However, even if the genetic makeup of rice hosts explained the largest proportion of variance in the composition of

Table 2. The proportion of core reads and core OTUs within and among the rice genetic groups.

Community	Plant compartment	Genotype group	Number of core reads	Total number of reads	Proportion of core reads	Number of core OTUs	Total number of OTUs	Proportion of core OTUs
Bacterial	Stem	Modern + HHRTS landraces	18 703	101 850	18.3	1	91	1.0
		Modern	0	44 100	0.0	0	63	0.0
		HHRTS landraces	11 233	57 750	19.5	1	81	1.2
	Root	Modern + HHRTS landraces	41 838	120 666	34.7	3	325	0.9
		Modern	28 417	56 784	50.0	6	240	2.5
		HHRTS landraces	23 336	63 882	36.5	2	288	0.7
Fungal	Stem	Modern + HHRTS landraces	13 737	45 240	30.4	2	110	1.8
		Modern	8185	20 590	39.8	2	60	3.3
		HHRTS landraces	5501	24 650	22.3	2	91	2.2
	Root	Modern + HHRTS landraces	6534	30 858	21.2	2	105	1.9
		Modern	2729	13 344	20.4	1	71	1.4
		HHRTS landraces	1898	17 514	10.8	1	79	1.3

microbial communities in Malizhai HHRTS (15%–22%), other factors not covered by our study may also explain the observed patterns of microbial diversity. Indeed, the small sampling area selected (<2 km²) was more environmentally heterogeneous than initially envisioned. It is known that plant root microbial communities are influenced by soil cultivation histories and agricultural practices (Peiffer *et al.*, 2013; Li *et al.*, 2019). Hence, the turnover of rice varieties in HHRTS fields and other characteristics of the flooded terraces—such as the use of living animal fertilizers (ducks, snails, fish and cattle)—may also explain the heterogeneity of the microbial communities of different rice plants (Xie *et al.*, 2011; Jiao *et al.*, 2012). Precise information pertaining to the soil compositions or fertilizer usage history of our sampling sites is not available. However, previous surveys of farming practices in the portion of the HHRTS that we studied revealed that farmers employ a variety of cropping practices and frequently alter the diverse set of rice varieties that they use (Dedeurwaerdere and Hannachi, 2019).

Another factor that could have contributed to environmental heterogeneity is the presence of rice pathogens. Several recent studies have indicated both that the impact of microbial communities on the course and outcome of host diseases can be substantial, and that host diseases can reciprocally have a large impact on microbial community compositions (Berendsen *et al.*, 2012; Ritpitakphong *et al.*, 2016; Koskella *et al.*, 2017; Zhang *et al.*, 2018; Vannier *et al.*, 2019). We have recently

shown that the SRBSDV was prevalent in 2016 in the Malizhai HHRTS and that 23 of the 166 rice plants examined in the present study were infected by SRBSDV (Alonso *et al.*, 2019). We, therefore, examined the impact of SRBSDV infections on the compositions of microbial communities but no significant difference was observed between the microbial communities of SRBSDV-infected rice plants and those of SRBSDV-uninfected plants (Table S9). Interestingly, several of the HHRTS landraces and modern rice varieties had no SRBSDV infected plants. If these landraces and varieties are either resistant or tolerant to SRBSDV (Alonso *et al.*, 2019), it is most probable that this resistance would be attributable to host genetic factors, i.e. host resistance, rather than to their respective microbial communities. SRBSDV is transmitted by a flying insect such that root microbial communities are less likely to play a significant role in SRBSDV transmission than in the transmission of soil-borne rice pathogens.

Finally, in addition to deterministic factors, like host genotype that did explain the largest proportion of variance in the compositions of rice microbial communities, and other factors not covered by this study, we cannot rule out that stochastic processes could have also influenced microbial community compositions. Indeed, stochastic processes might account for the high variability in microbial community compositions that were observed both between plants belonging to the same rice genotypes and between plants sampled from the same HHRTS paddy fields (Zhou

and Ning, 2017). Stochasticity could for instance arise during the colonization of plants by microbes wherever early arriving taxa modify the surface or within-plant niches, making these niches more or less suitable for later-arriving species (Maignien *et al.*, 2014). Therefore, besides better disentangling the roles that biotic and abiotic factors play in modifying the microbial community of HHRTS rice plants, further studies will be needed to determine the impacts over time of stochastic processes on dynamic changes in rice microbial community compositions.

Materials and methods

Study area and sampling site

The village of Malizhai is located near the town of Xinjie in the Honghe Hani Yi Autonomous Prefecture (Yunnan province, China). This village has recently adopted a mixed landrace/modern variety system but, prior to 2010 was representative of the cultural landscape of the Hani rice terraces (Jiao *et al.*, 2012). The sampling site in Malizhai covered an area of approximately 2 km² (N23°07'55.04" E102°46'03.95") at an altitude ranging between 1570 and 1608 m.

Plant sampling

Nineteen small Malizhai paddy fields were sampled in July 2016, including 11 fields cultivated with HHRTS landraces and eight fields cultivated with "modern rice varieties". Ten plants were collected per field, regardless of the presence of disease symptoms. It is noteworthy that at the time of the sampling survey disease pressure was very low and a very few symptoms were visible on rice plants throughout the Malizhai terraces. Roots, stems and leaves were separately collected and rinsed with water in order to remove the attached soil. All samples were individually stored at 4°C in a mobile fridge. Within 24 h, all samples were dried in the presence of CaCl₂ until DNA extraction.

Plant and microbial DNA extractions

Plant DNA extractions and genotyping-by-sequencing GBS were carried out on 30 mg of dried individual leaf samples as described previously (Alonso *et al.*, 2019). We performed microbial DNA extractions on each stem sample with both epiphyte and endophyte microorganisms and on each root sample with its rhizoplane (root surface) and endophyte microorganisms. In total, 30 mg of the 190 sample rice stems and the 190 sample rice roots were individually frozen in liquid nitrogen then ground with bead beating (two steel beads, diameter 0.1 cm and one ceramic bead, diameter 0.5 cm) using a FastPrep-24 5G

System (MP Biomedicals - Fisher Scientific) for 2 × 30 s at 6 ms⁻¹. Total genomic DNA was extracted from the resulting powder using the NucleoMag Plant Kit (Macherey–Nagel, Germany) and KingFisher Flex Purification System (ThermoFisher Scientific, MA, USA), following the manufacturer's instructions.

Analysis of rice genotyping-by-sequencing data

GBS reads for rice landraces and whole-genome re-sequencing data for nine randomly selected individuals from four representative subpopulations of *indica* (XI-1A, XI-1B, XI-2 and XI-3) described in the 3000 Rice Genomes Project (Wang *et al.*, 2018) were aligned against the Nipponbare rice reference genome (MSU7) using Bowtie2.3.5 (Langmead *et al.*, 2009; Langmead and Salzberg, 2012; Wang *et al.*, 2018; Langmead *et al.*, 2019). The GBS raw sequence read data are available from the NCBI Sequence Read Archive, under the BioProject ID number PRJNA573048. SNP-calling using bcftools mpileup (options --max-depth 500 -a DP) was carried out independently for the 3000 Rice Genomes Project dataset and the GBS dataset (<https://www.sanger.ac.uk/science/tools/samtools-bcftools-htslib>). For the GBS dataset, sites with either AN ≤ 300 or MQ ≤ 20 or which were tagged as 'LowQual' were removed using bcftools filter (INFO/AN ≤ 300; INFO/MQ ≤ 20) and bash (grep -v 'LowQual'). Genotypes with DP ≤ 5 were converted to missing data using vcftools (--minDP 5). For the 3000 Rice Genomes Project dataset, sites with DP ≤ 5 or MQ ≤ 20 were filtered out using bcftools filter (INFO/DP ≤ 5; INFO/MQ ≤ 20). The VCF files for the 3000 Rice Genomes Project and GBS accessions were merged using bcftools merge. The merged dataset included 143 611 sites with ≤50% missing data and 10 028 sites with ≤10% missing data. We used the dataset with ≤10% missing data for analyses of genealogical relationships among genotypes.

We constructed a neighbour-network using SplitsTree 4.13, to visualize evolutionary relationships between the *indica* rice genotypes while taking the possibility of recombination or incomplete lineage sorting into account (Huson and Bryant, 2006). We used the program sNMF (<http://membres-timc.imag.fr/Olivier.Francois/snmf/index.htm>) to test the dataset with ≤50% missing data for evidence of population subdivision by partitioning genotypes into K ancestral populations and estimating individual ancestry coefficients in the K populations (Frichot *et al.*, 2014). We ran sNMF for a number of clusters K-values ranging from 1 to 15, and for each K we performed 10 replicates. We used the Greedy algorithm in CLUMPP 1.1.2 (<http://web.stanford.edu/group/rosenberglab/clumpp.html>) (Jakobsson and Rosenberg, 2007) to identify runs representing the same clustering solution (i.e. same mode). We randomly selected one representative of runs belonging to the major

mode for graphical representation as a stacked barplot using the Matplotlib package in Python. Nucleotide diversity (π) was estimated using the scikit-allel in Python (<https://github.com/cggh/scikit-allel>).

PCR amplification and sequencing

The composition and diversity of rice-associated microbial communities were characterized by applying a high-throughput sequencing-based protocol that targets PCR-generated amplicons. Bacterial communities were characterized from the variable region, V3-V4, of the 16S rRNA gene using the primers 341-F (5'-CTACGGGNGGCWGCAG 3') and 785-R (5'-GACTACHVGGGTATCTAATCC3') as universal primers to maximize bacterial taxonomic assignment (Thijs *et al.*, 2017). We used ITS86-F (5'-GTGAATCAT CGAATCTTTGAA3') and ITS4-R (5'-TCCTCCGCTTATTG ATATGC3') as universal primers for amplification of the fungal ITS1 region (Op De Beeck *et al.*, 2014).

DNA amplification was performed by PCR in a total volume of 25 μ l containing 1 $\times$ GoTaq G2 DNA polymerase buffer (Promega Corporation, Madison, USA), 0.5 μ M of each primer, and 0.2 μ M dNTPs and 1 μ l of genomic DNA. We used peptide nucleotide acid (PNA) clamps, which specifically bind to mitochondrial or chloroplast sequences to block the amplification of plant derived-DNA. We added to the PCR mix PNA blocker oligos (PNA Bio, Thousand Oaks, CA, USA) at 0.5 μ M targeting the 16S rRNA gene of plant mitochondria (PNAm: GGCAAGTGTCTTCGGA), and chloroplasts (PNSp: GGCTCAACCCTGGACAG) (Jackrel *et al.*, 2017). All amplifications were performed in a thermal cycler (Biometra, Gottingen, Germany) under the following conditions for 16S rRNA gene amplifications: an initial denaturation at 98°C for 3 min followed by 30 cycles of denaturation at 98°C for 15 s, PNA annealing at 75°C for 10 s, primers annealing at 52°C for 10 s, extension at 72°C for 30 s, and a final extension at 72°C for 10 min. ITS amplifications were performed under the following conditions: an initial denaturation at 94°C for 3 min followed by 30 cycles of denaturation at 94°C for 45 s, primers annealing at 55°C for 45 s, extension at 72°C for 2 min, and a final extension at 72°C for 10 min. Each PCR product was tagged with a combination of two different barcodes designed by a genomic platform (GenSeq, University of Sciences, UMII, Montpellier, France) that allows for the identification of 384 different PCR products loaded onto the same MiSeq flow cell. Negative controls from the extraction step and PCR reaction were sequenced with the plant samples to evaluate and exclude contaminant reads from the sample data set. All PCR products were pooled and purified, and the library was constructed and sequenced using a GenSeq platform with Illumina paired-end 2 $\times$ 250-bp technology and V2 chemistry.

Sequence processing, OTU clustering, and OTU filtering

Base calling and demultiplexing of Illumina sequences were carried out using RTA v1.18.54, MCS 2.6 and bcl2fastq2.17. Paired Illumina MiSeq reads were assembled with VSEARCH v2.11.0 (Rognes *et al.*, 2016) using the command `fastq_mergepairs` and the option `fastq_allowmergestagger`. Primer clipping was performed with cutadapt v1.9 (Martin, 2011) allowing a 2/3-length partial match for forward and reverse primers. Only reads containing both primers were retained. The expected error per read was estimated with the VSEARCH command `fastq_filter` and the option `eeout`. Each sample was then dereplicated by merging identical reads using vsearch's command, `derep_fulllength`, and converted to FASTA format. To prepare for clustering, the samples were pooled and further dereplicated with VSEARCH. Files containing per-read expected error values were also dereplicated to retain only the lowest expected error for each unique sequence. Clustering was performed with Swarm v3.0.0 (Mahe *et al.*, 2015), using a local threshold of one difference and the `fastidious` option.

OTU representative sequences were then searched for evidence of chimeras with the VSEARCH command, `uchime_denovo` (Edgar *et al.*, 2011). In parallel, representative sequences were assigned to taxa using the stampa pipeline (<https://github.com/frederic-mahe/stampa/>) and the ribosomal database SILVA v138 (<https://www.arb-silva.de/>) (Quast *et al.*, 2013) for the bacterial community, and a custom version of the ITS database UNITE v8 (<https://unite.ut.ee/>) (Abarenkov *et al.*, 2010; Koljalg *et al.*, 2013) for the fungal community.

Clustering results, expected error values, taxonomic assignments and chimera detection results were used to build a raw OTU table. Up to that point, reads without primers, reads shorter than 32 nucleotides and reads with uncalled bases ('N') had been eliminated. To create the 'cleaned' OTU table, additional filters were applied to retain only non-chimeric OTUs, OTUs with an expected error per nucleotide below 0.0002, OTUs containing more than three reads or which were found two or more samples. ITS and 16S OTUs tables obtained after the initial cleaning step (chimeric OTUs removal) can be found at https://github.com/P-alonso/HHRTS_microbial_diversit (Yunnan_Rice_2016_16S_roots_and_stems_384_samples.OTU.filtered.table, Yunnan_Rice_2016_ITS2_roots_and_stems_384_samples.OTU.filtered.table). All 16S and ITS sequences with a higher abundance in at least one negative control than the rice samples were excluded from the final dataset. All 16S OTUs assigned to chloroplast or mitochondrial sequences were excluded. Similarly, ITS OTUs not assigned to fungal reference sequences were excluded. All codes and representative OTU sequences can be found in HTML format in Supporting information file S1. The raw data are available

from the NCBI Sequence Read Archive (SRA) under BioProject ID number PRJNA573048.

Statistical analyses of microbial community data

All statistical analyses were performed in R (<http://www.R-project.org>) with the Phyloseq (McMurdie and Holmes, 2013), Vegan (Oksanen *et al.*, 2018), pairwiseAdonis (Martinez Arbizu, 2020) and Metacoder packages. Biological replicates corresponding to 10 plants per field were used to fix a biological reproducibility threshold in order to perform the sequencing de-noising. Only OTUs with over 50% prevalence per paddy field were considered. Samples with less than 1000 reads for bacterial communities and less than 5000 reads for fungal communities were discarded following which the OTUs abundances were rarefied to homogenize sequencing depth.

We estimated the microbial diversity using richness and Shannon's diversity indices for α -diversity calculations and UniFrac distances for β -diversity calculations. To assess the relationships between plant genotypes (HHRTS landraces vs. modern rice varieties) and plant-associated microbial community, a Tukey HSD test and a PERMANOVA were performed for the α -diversity and the β -diversity respectively. Differences in community composition between groups of samples were statistically evaluated by PERMANOVA using the UniFrac distance matrices. The `adonis()` function was used to calculate PERMANOVA with 10 000 permutations between HHRTS and modern varieties. A pairwise *post hoc* comparison was performed to evaluate the difference in community compositions across the six rice subgroups (HHRTS landraces subgroup1, HHRTS landraces subgroup2, HHRTS landraces subgroup3, HHRTS landraces subgroup4, modern subgroup1 and modern subgroup2) using the `pairwise.adonis()` function. Differences in community composition were assessed by PCoA based on weighted UniFrac and unweighted UniFrac distances. PCoA is an ordination method that represents pairwise (dis)similarities between samples in a low-dimensional space, so that samples placed closer in the graph are more similar than those placed further apart (McMurdie and Holmes, 2013). R codes used in statistical analyses are provided at https://github.com/P-alonso/HHRTS_microbial_diversity/.

The Metacoder package was used to visualize differential abundances in taxa between the modern varieties and HHRTS landrace groups using the function `compare_groups()` among plant microbial communities on a differential heat tree. The ratio of the mean OTU abundance between landraces and modern rice varieties was calculated. For each taxon, a Wilcoxon Rank Sum test was used to test for differences between the median abundances of samples between landraces and modern rice varieties. The taxon abundance of bacterial and

fungal communities was plotted on a taxonomic tree and the result of the Wilcoxon Rank Sum test with a fold-change cutoff of 1.5 and a *P*-value cutoff of 0.05 were used to highlight the differences in taxon abundance between HHRTS landraces and modern rice varieties. To test for significant associations between microbial community dissimilarities and the phylogenetic distances between rice plants that hosted these communities, we conducted partial Mantel tests, as implemented in the `vegan` package in R (Oksanen *et al.*, 2018) between the unweighted UniFrac distances of microbial communities and the patristic distances of rice plants calculated with the `cophenetic()` function.

The core microbial taxa within the stem and root microbial communities were identified using the Microbiome R package based on a criterion of prevalence in at least 80% of the samples from the 18 *indica* varieties (including all the modern varieties and HHRTS landraces) with no criterion related to the relative abundance of the taxa, in order to consider rare but prevalent microbial taxa. A second analysis focused on stem and root microbial taxa identified from each group of rice genotypes using the same criterion. Based on this criterion, a list of core taxa was identified and their relative abundance was calculated.

Acknowledgements

This study was funded by the French International and Agricultural Research Agency (CIRAD), the Agropolis Fondation (E-Space Flagship Program, grant number 1504-004), CGIAR Research Program (CRP) on rice agri-food systems (RICE, 2017-2022) and the French National Research Agency (Next Generation Biomonitoring Project, grant number ANR-17-CE32-0011). We thank the Yunnan Agricultural University and the International Associated Laboratory Plantomix (INRAE/Yunnan Agricultural University) for their technical logistics and support. D.P.M. has received a research grant from the National Research Foundation of South Africa.

References

- Abarenkov, K., Henrik Nilsson, R., Larsson, K.-H., Alexander, I.J., Eberhardt, U., Erland, S., *et al.* (2010) The UNITE database for molecular identification of fungi – recent updates and future perspectives. *New Phytol* **186**: 281–285.
- Alonso, P., Gladieux, P., Moubset, O., Shih, P.J., Mournet, P., Frouin, J., *et al.* (2019) Emergence of southern rice black-streaked dwarf virus in the centuries-old Chinese Yuanyang Agrosystem of rice landraces. *Viruses* **11**: v11110985.
- Berendsen, R.L., Pieterse, C.M., and Bakker, P.A. (2012) The rhizosphere microbiome and plant health. *Trends Plant Sci* **17**: 478–486.
- Bernardo, P., Charles-Dominique, T., Barakat, M., Ortet, P., Fernandez, E., Filloux, D., *et al.* (2018) Geometagenomics

- illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *ISME J* **12**: 173–184.
- Bouffaud, M.L., Poirier, M.A., Muller, D., and Moenne-Loccoz, Y. (2014) Root microbiome relates to plant host evolution in maize and other *Poaceae*. *Environ Microbiol* **16**: 2804–2814.
- Bulgarelli, D., Garrido-Oter, R., Munch, P.C., Weiman, A., Droge, J., Pan, Y., et al. (2015) Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* **17**: 392–403.
- Bulgarelli, D., Schlaeppi, K., Spaepen, S., Ver Loren van Themaat, E., and Schulze-Lefert, P. (2013) Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* **64**: 807–838.
- Cardinale, M., Grube, M., Erlacher, A., Quehenberger, J., and Berg, G. (2015) Bacterial networks and co-occurrence relationships in the lettuce root microbiota. *Environ Microbiol* **17**: 239–252.
- Coleman-Derr, D., Desgarennes, D., Fonseca-Garcia, C., Gross, S., Clingenpeel, S., Woyke, T., et al. (2016) Plant compartment and biogeography affect microbiome composition in cultivated and native *Agave* species. *New Phytol* **209**: 798–811.
- Compant, S., Samad, A., Faist, H., and Sessitsch, A. (2019) A review on the plant microbiome: ecology, functions, and emerging trends in microbial application. *J Adv Res* **19**: 29–37.
- Cordovez, V., Dini-Andreote, F., Carrion, V.J., and Raaijmakers, J.M. (2019) Ecology and evolution of plant microbiomes. *Annu Rev Microbiol* **73**: 69–88.
- Cui, B., You, Z., and Yao, M. (2008) Vertical characteristics of the Hani terraces paddyfield ecosystem in Yunnan, China. *Front Biol China* **3**: 351–359.
- Dedeurwaerdere, T., and Hannachi, M. (2019) Socio-economic drivers of coexistence of landraces and modern crop varieties in agro-biodiversity rich Yunnan rice fields. *Ecol Econ* **159**: 177–188.
- Ding, L.J., Cui, H.L., Nie, S.A., Long, X.E., Duan, G.L., and Zhu, Y.G. (2019) Microbiomes inhabiting rice roots and rhizosphere. *FEMS Microbiol Ecol* **95**: fuz040.
- Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200.
- Edwards, J., Johnson, C., Santos-Medellin, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., et al. (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci U S A* **112**: E911–E920.
- Edwards, J.A., Santos-Medellin, C.M., Liechty, Z.S., Nguyen, B., Lurie, E., Eason, S., et al. (2018) Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. *PLoS Biol* **16**: e2003862.
- Emmett, B.D., Youngblut, N.D., Buckley, D.H., and Drinkwater, L.E. (2017) Plant phylogeny and life history shape rhizosphere bacterial microbiome of summer annuals in an agricultural field. *Front Microbiol* **8**: 2414.
- Fernández-Valiente, E., and Quesada, A. (2004) A shallow water ecosystem: rice-fields. The relevance of cyanobacteria in the ecosystem. *Limnetica* **23**: 95–107.
- Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., and Francois, O. (2014) Fast and efficient estimation of individual ancestry coefficients. *Genetics* **196**: 973–983.
- Hacquard, S., Spaepen, S., Garrido-Oter, R., and Schulze-Lefert, P. (2017) Interplay between innate immunity and the plant microbiota. *Annu Rev Phytopathol* **55**: 565–589.
- Hall, M.W., Singh, N., Ng, K.F., Lam, D.K., Goldberg, M.B., Tenenbaum, H.C., et al. (2017) Inter-personal diversity and temporal dynamics of dental, tongue, and salivary microbiota in the healthy oral cavity. *NPJ Biofilms Microbiomes* **3**: 2.
- Hamonts, K., Trivedi, P., Garg, A., Janitz, C., Grinyer, J., Holford, P., et al. (2018) Field study reveals core plant microbiota and relative importance of their drivers. *Environ Microbiol* **20**: 124–140.
- Hassani, M.A., Özkurt, E., Seybold, H., Dagan, T., and Stukenbrock, E.H. (2019) Interactions and coadaptation in plant metaorganisms. *Annu Rev Phytopathol* **57**: 483–503.
- Huson, D.H., and Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**: 254–267.
- Jackrel, S.L., Owens, S.M., Gilbert, J.A., and Pfister, C.A. (2017) Identifying the plant-associated microbiome across aquatic and terrestrial environments: the effects of amplification method on taxa discovery. *Mol Ecol Resour* **17**: 931–942.
- Jakobsson, M., and Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.
- Jiao, Y., Li, X., Liang, L., Takeuchi, K., Okuro, T., Zhang, D., and Sun, L. (2012) Indigenous ecological knowledge and natural resource management in the cultural landscape of China's Hani terraces. *Ecol Res* **27**: 247–263.
- Keesing, F., Holt, R.D., and Ostfeld, R.S. (2006) Effects of species diversity on disease risk. *Ecol Lett* **9**: 485–498.
- Kim, H., Lee, K.K., Jeon, J., Harris, W.A., and Lee, Y.-H. (2020) Domestication of *Oryza* species eco-evolutionarily shapes bacterial and fungal communities in rice seed. *Microbiome* **8**: 20.
- Knief, C., Delmotte, N., Chaffron, S., Stark, M., Innerebner, G., Wassmann, R., et al. (2012) Meta-proteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *ISME J* **6**: 1378–1390.
- Koljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F., Bahram, M., et al. (2013) Towards a unified paradigm for sequence-based identification of fungi. *Mol Ecol* **22**: 5271–5277.
- Koskella, B., Meaden, S., Crowther, W.J., Leimu, R., and Metcalf, C.J.E. (2017) A signature of tree health? Shifts in the microbiome and the ecological drivers of horse chestnut bleeding canker disease. *New Phytol* **215**: 737–746.
- Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.
- Langmead, B., Wilks, C., Antonescu, V., and Charles, R. (2019) Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* **35**: 421–432.

- Leff, J.W., Lynch, R.C., Kane, N.C., and Fierer, N. (2017) Plant domestication and the assembly of bacterial and fungal communities associated with strains of the common sunflower, *Helianthus annuus*. *New Phytol* **214**: 412–423.
- Li, X., Jousset, A., de Boer, W., Carrion, V.J., Zhang, T., Wang, X., and Kuramae, E.E. (2019) Legacy of land use history determines reprogramming of plant physiology by soil microbiome. *ISME J* **13**: 738–751.
- Liao, J., Huang, H., Meusnier, I., Adreit, H., Ducasse, A., Bonnot, F., et al. (2016) Pathogen effectors and plant immunity determine specialization of the blast fungus to rice subspecies. *Elife* **5**: e19377.
- Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., et al. (2012) Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**: 86–90.
- Mahe, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2015) Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* **3**: e1420.
- Maignien, L., DeForce, E.A., Chafee, M.E., Eren, A.M., and Simmons, S.L. (2014) Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* phyllosphere communities. *mBio* **5**: e00682-00613.
- Martin, M.P. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: 200.
- Martinez Arbizu, P. (2020) pairwiseAdonis: Pairwise multi-level comparison using adonis. R package version 0.4.
- McMurdie, P.J., and Holmes, S. (2013) Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**: e61217.
- Meyer, R.S., and Purugganan, M.D. (2013) Evolution of crop species: genetics of domestication and diversification. *Nat Rev Genet* **14**: 840–852.
- Moronta-Barrios, F., Gionechetti, F., Pallavicini, A., Marys, E., and Venturi, V. (2018) Bacterial microbiota of rice roots: 16S-based taxonomic profiling of endophytic and rhizospheric diversity, endophytes isolation and simplified endophytic community. *Microorganisms* **6**: 6010014.
- Mutch, L.A., and Young, J.P. (2004) Diversity and specificity of *Rhizobium leguminosarum* biovar *viciae* on wild and cultivated legumes. *Mol Ecol* **13**: 2435–2444.
- Ofek-Lalzar, M., Sela, N., Goldman-Voronov, M., Green, S. J., Hadar, Y., and Minz, D. (2014) Niche and host-associated functional signatures of the root surface microbiome. *Nat Commun* **5**: 4950.
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D. et al. (2018) Vegan: Community Ecology Package. R package version 2.5.3.
- Op De Beeck, M., Lievens, B., Busschaert, P., Declerck, S., Vangronsveld, J., and Colpaert, J.V. (2014) Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. *PLoS One* **9**: e97629.
- Peiffer, J.A., and Ley, R.E. (2013) Exploring the maize rhizosphere microbiome in the field: a glimpse into a highly complex system. *Commun Integr Biol* **6**: e25177.
- Peiffer, J.A., Spor, A., Koren, O., Jin, Z., Tringe, S.G., Dangl, J.L., et al. (2013) Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci U S A* **110**: 6548–6553.
- Perez-Jaramillo, J.E., Carrion, V.J., de Hollander, M., and Raaijmakers, J.M. (2018) The wild side of plant microbiomes. *Microbiome* **6**: 143.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., and Yarza, P. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.
- Redford, A.J., Bowers, R.M., Knight, R., Linhart, Y., and Fierer, N. (2010) The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ Microbiol* **12**: 2885–2893.
- Reinhold-Hurek, B., Bonger, W., Burbano, C.S., Sabale, M., and Hurek, T. (2015) Roots shaping their microbiome: global hotspots for microbial activity. *Annu Rev Phytopathol* **53**: 403–424. <https://doi.org/10.1146/annurev-phyto-082712-102342>.
- Ritpitakphong, U., Falquet, L., Vimoltust, A., Berger, A., Metraux, J.P., and L'Haridon, F. (2016) The microbiome of the leaf surface of *Arabidopsis* protects against a fungal pathogen. *New Phytol* **210**: 1033–1043.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahe, F. (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**: e2584.
- Roossinck, M.J., and Garcia-Arenal, F. (2015) Ecosystem simplification, biodiversity loss and plant virus emergence. *Curr Opin Virol* **10**: 56–62.
- Sapkota, R., Knorr, K., Jorgensen, L.N., O'Hanlon, K.A., and Nicolaisen, M. (2015) Host genotype is an important determinant of the cereal phyllosphere mycobiome. *New Phytol* **207**: 1134–1144.
- Shenton, M., Iwamoto, C., Kurata, N., and Ikeo, K. (2016) Effect of wild and cultivated rice genotypes on rhizosphere bacterial community composition. *Rice* **9**: 42.
- Shi, S., Chang, J., Tian, L., Nasir, F., Ji, L., Li, X., and Tian, C. (2019) Comparative analysis of the rhizomicrobiome of the wild versus cultivated crop: insights from rice and soybean. *Arch Microbiol* **201**: 879–888.
- Stuckenbrock, E.H., and McDonald, B.A. (2008) The origins of plant pathogens in agro-ecosystems. *Annu Rev Phytopathol* **46**: 75–100.
- Szoboszlay, M., Lambers, J., Chappell, J., Kupper, J.V., Moe, L.A., and McNear, D.H. (2015) Comparison of root system architecture and rhizosphere microbial communities of Balsas teosinte and domesticated corn cultivars. *Soil Biol Biochem* **80**: 34–44.
- Theis, K.R., Dheilly, N.M., Klassen, J.L., Brucker, R.M., Baines, J.F., Bosch, T.C., et al. (2016) Getting the hologenome concept right: an eco-evolutionary framework for hosts and their microbiomes. *mSystems* **1**: 00028-00016.
- Thijs, S., Op De Beeck, M., Beckers, B., Truyens, S., Stevens, V., Van Hamme, J.D., et al. (2017) Comparative evaluation of four bacteria-specific primer pairs for 16S rRNA gene surveys. *Front Microbiol* **8**: 494.
- UNESCO. (2013) Convention concerning the protection of the world cultural and natural heritage. In *WHC-13/37COM/20*. Phnom Penh, Cambodia: UNESCO, p. 248.
- Vandenkoornhuys, P., Quaiser, A., Duhamel, M., Le Van, A., and Dufresne, A. (2015) The importance of the microbiome of the plant holobiont. *New Phytol* **206**: 1196–1206.
- Vannier, N., Agler, M., and Hacquard, S. (2019) Microbiota-mediated disease resistance in plants. *PLoS Pathog* **15**: e1007740.

- Verma, D., Garg, P.K., and Dubey, A.K. (2018) Insights into the human oral microbiome. *Arch Microbiol* **200**: 525–540.
- Vorholt, J.A. (2012) Microbial life in the phyllosphere. *Nat Rev Microbiol* **10**: 828–840.
- Wagner, M.R., Lundberg, D.S., Del Rio, T.G., Tringe, S.G., Dangl, J.L., and Mitchell-Olds, T. (2016) Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat Commun* **7**: 12151.
- Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., et al. (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* **557**: 43–49.
- Wang, W., Zhai, Y., Cao, L., Tan, H., and Zhang, R. (2016) Endophytic bacterial and fungal microbiota in sprouts, roots and stems of rice (*Oryza sativa* L.). *Microbiol Res* **188–189**: 1–8.
- Xie, J., Hu, L., Tang, J., Wu, X., Li, N., Yuan, Y., et al. (2011) Ecological mechanisms underlying the sustainability of the agricultural heritage rice-fish coculture system. *Proc Natl Acad Sci U S A* **108**: E1381–E1387.
- Yang, L., Liu, M., Lun, F., Yuan, Z., Zhang, Y., and Min, Q. (2017) An analysis on crops choice and its driving factors in agricultural heritage systems—a case of Honghe Hani Rice terraces system. *Sustainability* **9**: 1162.
- Zhang, Z., Luo, L., Tan, X., Kong, X., Yang, J., Wang, D., et al. (2018) Pumpkin powdery mildew disease severity influences the fungal diversity of the phyllosphere. *PeerJ* **6**: e4559.
- Zhou, J., and Ning, D. (2017) Stochastic community assembly: does it matter in microbial ecology? *Microbiol Mol Biol Rev* **81**: e00002–00017.
- Zhu, Y., Chen, H., Fan, J., Wang, Y., Li, Y., Chen, J., et al. (2000) Genetic diversity and disease control in rice. *Nature* **406**: 718–722.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

File S1 Codes and bioinformatics methods used to produce OTU tables after the initial cleaning (chimeric OTUs removal) (HTML format).

Fig. S1 Relationship between minimal cross-entropy and number of ancestral populations (K) modelled in the sNMF analysis of population subdivision.

Fig. S2 Violin plots of rice microbial alpha diversity (richness and Shannon diversity indexes) of (A) root and stem bacterial communities and (B) roots and stems fungal communities. P-values of Tukey HSD tests are shown with *** $P < 0.001$ or not shown if $P > 0.05$. PCoA plots based on unweighted UniFrac distances of (C) root and stem bacterial communities and (D) root and stem fungal communities. Communities identified from stems and roots are coloured in blue and red respectively. Axes represent the two dimensions explaining the greatest proportion of variances in the communities for each analysis. Permutational multivariate analysis of variance (PERMANOVA) results are indicated (R^2 and the P-value).

Fig. S3 Violin plots of rice microbial α diversity (richness and Shannon diversity indices) across the six subgroups of rice

sampled in the HHRTS (HHRTS landraces subgroup1 (L1), HHRTS landraces subgroup2 (L2), HHRTS landraces subgroup3 (L3), HHRTS landraces subgroup4 (L4), modern subgroup1 (M1) and modern subgroup2 (M2)) for (A) the rice stem bacterial communities, (B) the rice stem fungal communities, (C) the rice root bacterial communities, and (D) the rice root fungal communities. Results from Tukey HSD tests are presented as letters denoting groups that are significantly different (P-values < 0.05).

Fig. S4 PCoA plots based on weighted UniFrac distances of (A) rice stem bacterial communities and (C) rice stem fungal communities as well as (B) rice roots bacterial communities and (D) rice roots fungal communities. Communities from the HHRTS landraces and from the modern rice varieties are labelled in green and brown respectively. Axes represent the two dimensions explaining the greatest proportion of variances in the communities for each analysis. Results of a permutational multivariate analysis of variance (PERMANOVA) are indicated (R^2 and the P-value).

Fig. S5 Pairwise comparison of number of reads of rice stem bacterial communities assigned to bacterial species among the six HHRTS rice subgroups. The node width indicates the number of reads assigned to each taxonomic rank level along the tree and the colour indicates the statistically significant differences in relative bacterial taxa abundance. A taxon coloured brown is more abundant in the communities in the column and a taxon coloured green is more abundant in the communities in the row.

Fig. S6 Pairwise comparison of number of reads of rice root bacterial communities assigned to bacterial species among the six HHRTS rice subgroups. The node width indicates the number of reads assigned to each taxonomic rank level along the tree and the colour indicates the statistically significant differences in relative bacterial taxa abundance. A taxon coloured brown is more abundant in the communities in the column and a taxon coloured green is more abundant in the communities in the row.

Fig. S7 Pairwise comparison of number of reads of rice stem fungal communities assigned to fungal species among the six HHRTS rice subgroups. The node width indicates the number of reads assigned to each taxonomic rank level along the tree and the colour indicates the statistically significant differences in relative fungal taxa abundance. A taxon coloured brown is more abundant in the communities in the column and a taxon coloured green is more abundant in the communities in the row.

Fig. S8 Pairwise comparison of number of reads of rice root fungal communities assigned to fungal species among the six HHRTS rice subgroups. The node width indicates the number of reads assigned to each taxonomic rank level along the tree and the colour indicates the statistically significant differences in relative fungal taxa abundance. A taxon coloured brown is more abundant in the communities in the column and a taxon coloured green is more abundant in the communities in the row.

Table S1 Abundance of reads and OTUs in bacterial and fungal community data sets for stem and root samples through the different steps of the bio-informatics treatments.

Table S2 Abundance of bacterial OTUs in the rice communities of modern varieties (YYM) and HHRTS landraces (YYT)

obtained from root (R as last letter) or stem (T as last letter) and taxonomic assignation. Different numbers indicate different rice paddies. Cleaned data obtained after the rarefaction process.

Table S3 Abundance of fungal OTUs in the rice communities of modern varieties (YYM) and HHRTS landraces (YYT) obtained from root (R as last letter) or stem (T as last letter) and taxonomic assignation. Different numbers indicate different rice paddies. Cleaned data obtained after the rarefaction process.

Table S4 Pairwise permutational multivariate analysis of variance (PERMANOVA) results for bacterial and fungal communities for both stem and root samples using unweighted UniFrac distance values (10,000 permutations), R^2 denotes the proportion of variance that could be explained by the grouping. P. adjusted corresponds to the Bonferroni correction applied to adjust the P-value for multiple comparisons within each group. Sig significance level *P. adjusted <0.05, **P. adjusted <0.01 and NS not significant.

Table S5 Bacterial taxa from stems and roots that are significantly differentially abundant between the six rice subgroups

using Wilcoxon signed rank tests. Only, significant P-values <0.05 are indicated.

Table S6 Fungal taxa from stems and roots that are significantly differentially abundant between the six rice subgroups using Wilcoxon signed rank tests. Only, significant P-values <0.05 are indicated.

Table S7 Detection of phylogenetic signal in root-associated and stem-associated microbial communities. Mantel statistic based on Pearson's product-moment correlation to assess the correlation between unweighted UniFrac distances matrix of microbial community dissimilarities and the rice genetic distances (10,000 permutations), R^2 denotes the proportions of variances that could be explained by the grouping.

Table S8 List of the taxa that were present in more than 80% of the plant samples of each rice genetic group.

Table S9 Permutational multivariate analysis of variance (PERMANOVA) results for the influence of infection of SRBSDV on bacterial and fungal composition using UniFrac distance values (10,000 permutations), R^2 denotes the proportions of variances that could be explained by the grouping.

2.5. Conclusions du chapitre 2

L'objectif initial de ce chapitre de thèse était de vérifier si l'introduction de variétés modernes dans l'agrosystème du Yuanyang pouvait avoir un impact sur la composition du microbiote associé au riz. Les trois parties de ce travail ont permis d'obtenir sept conclusions majeures :

1. Les variétés de riz cultivés dans l'agrosystème du Yuanyang sont très diversifiées. L'introduction des variétés modernes entraîne un changement de la composition en génotypes de riz cultivés dans cet agrosystème, mais ne réduit que faiblement la diversité génétique globale des riz cultivés dans ce dernier.
2. Le profil génétique de la variété de riz Hongyang, jusqu'à lors présentée comme une variété moderne dans une étude sociologique (Dedeurwaerdere & Hannachi, 2019), a permis de montrer que cette dernière est génétiquement proche des variétés traditionnelles de l'agrosystème du Yuanyang. La variété Hongyang n'a pas de lien génétique avec les autres variétés modernes cultivées introduites dans cet agrosystème.
3. Le riz cultivé dans l'agrosystème du Yuanyang présente un virome essentiellement constitué de mycovirus et d'un virus pathogène du riz le : Southern rice black-streaked dwarf virus (SRBSDV).
4. L'épidémie de SRBSDV dans le bassin versant de la rivière Malizhai a été confirmée sur deux années (2016 et 2018).
5. Le génotype du riz impacte fortement la prévalence virale du SRBSDV et n'a qu'un effet très modéré sur la structure des communautés bactériennes et fongiques associées au riz.
6. L'infection de plants de riz par le SRBSDV n'a pas d'effet significatif sur la structure des communautés bactériennes et fongiques associées au riz.
7. Le microbiote associé au riz dans l'agrosystème du Yuanyang est hautement diversifié. Le microbiote cœur du riz dans l'agrosystème du Yuanyang représente seulement moins de 2,5% des taxa.

Chapitre 3: Etude de facteurs de structuration du microbiote endophyte du riz sur un site d'étude en Camargue

3.1. Introduction

L'étude bibliographique générale présentée au début de ce manuscrit, souligne l'importance de caractériser la structure des communautés microbiennes associées à la plante. Elle permet de mieux comprendre les interactions plante/microorganismes qui participent au bon fonctionnement de la plante, ainsi qu'à son adaptation à des stress biotiques et abiotiques. Les travaux effectués dans le chapitre 2, ont mis en évidence des importantes différences de composition du microbiote entre différents organes de riz échantillonnés en fin de tallage. Des différences marginales, mais significatives, ont été observées entre les différents génotypes de riz cultivés dans un même agrosystème, à un temps donné. Nous avons alors suggéré qu'une part majeure de la variabilité de la structure du microbiote du riz à la fin du stade de développement tallage dans les terrasses du Yuanyang, est très probablement expliquée par des facteurs autres que le génotype de la plante (Hartman et al., 2018, Silva et al., 2017, Nemergut et al., 2013). Cette hypothèse reflète que la diversité du microbiote du riz, observée dans les terrasses chinoises, résulte potentiellement d'une grande diversité des techniques et historiques culturels de chaque microparcelle, et possiblement de caractéristiques du sol variables entre les parcelles. Malheureusement, ce dispositif expérimental « paysan » ne nous a clairement pas permis de maîtriser certaines conditions culturelles (rotation culturale, élevage animal dans les parcelles, utilisation d'engrais, de pesticides, etc.), du fait que les parcelles étaient gérées par des agriculteurs différents et qu'aucune donnée sur ce sujet n'a pu être collectée. Par ailleurs, en raison de la distance du lieu d'échantillonnage, nous n'avons caractérisé le microbiote du riz qu'à un temps donné et sur des échantillons séchés qui ne permettent pas de faire une dissociation de la flore épiphyte et de la flore endophyte. Or, il a été montré dans de précédentes études qu'il existe à la fois une dynamique temporelle dans la structuration des communautés microbiennes du riz (Zhang et al., 2018, Edwards et al., 2018), et qu'il est intéressant de dissocier les communautés endophytes, car elles peuvent être potentiellement impliquées dans des phénomènes de résistance à des maladies (Santhanam et al., 2015, Compant et al., 2019, Brader et al., 2017, Busby et al., 2016). De plus, ces communautés endophytes se distinguent des communautés extérieures à la plante et doivent subir des contraintes fortes de la part de la plante et notamment, être adaptées à son système immunitaire (Hardoim et al., 2008). Nous émettons l'hypothèse que ces communautés endophytes sont plus dépendantes du génotype de l'hôte. Enfin, nous n'avons pas pu étudier l'impact du microbiote du riz sur les maladies fongiques et bactériennes, et réciproquement, au sein de l'agrosystème des terrasses du Yuanyang, car aucune épidémie notable n'était présente. Nous avons cependant pu de façon

opportuniste, suite à l'analyse de métagénomique virale et à la caractérisation du Southern rice black-streaked dwarf virus, tester l'effet de la composition du microbiote sur l'aspect qualitatif (plante saine vs. plante malade) du pouvoir pathogène de ce virus du riz. Cette analyse n'a pas révélé d'effet significatif du microbiote sur l'infection du virus (et *vice-versa*).

Ainsi, pour mieux appréhender l'effet du génotype dans la dynamique de l'assemblage du microbiote, nous avons mis en place dans cette deuxième partie de la thèse, une étude expérimentale sur dispositifs expérimentaux *ad hoc* de terrain. Cette approche expérimentale de terrain a permis de mieux contrôler les étapes culturales et d'homogénéiser les variations environnementales pour mieux tester l'effet du génotype du riz sur la structuration du microbiote endophyte, et de pouvoir en suivre la dynamique dans le temps et dans l'espace. Ce dispositif avait aussi initialement pour objectif d'essayer de tester l'effet de la pyriculariose, qui est la principale maladie fongique du riz à l'échelle mondiale, sur le microbiote endophyte du riz, et réciproquement.

Pour se faire, nous avons eu l'opportunité de nous associer au projet de recherche pluridisciplinaire « SElection pour la résistance à la PYriculariose et l'utilisation de l'Azote (SEPYA) », mis en place par le Centre Français du Riz (CFR), en collaboration avec deux unités de recherches (l'UMR Biologie et génétique des interactions plante-parasite et l'UMR Amélioration Génétique et Adaptation des Plantes). Ce dispositif expérimental de terrain correspondait, en effet, à nos besoins en termes de bonne maîtrise des conditions culturales, d'utilisation de nombreux génotypes, et de sa proximité géographique. Ce projet avait pour objectif principal, d'identifier des variétés de riz tempérées combinant une bonne résistance à la pyriculariose et une bonne capacité d'utilisation de l'azote afin de limiter les apports d'intrants. En nous associant à cet essai, nous avons la possibilité de suivre les variations de compositions du microbiote de la partie aérienne (endosphère de la tige) et de la partie souterraine (endosphère de la racine) séparément pour plusieurs génotypes de riz au cours de leurs cycles de croissance sous pression d'un champignon phytopathogène (*Pyricularia oryzae*).

Les objectifs spécifiques de nos travaux étaient i) d'étudier les variations de richesse spécifique et de structure du microbiote endophyte du riz, au cours de son cycle de développement, ii) de tester l'effet du génotype sur la structuration du microbiote, à différentes étapes du développement du riz et iii) d'étudier l'effet de la pyriculariose sur le microbiote endophyte (et *vice versa*) de variétés de riz résistantes et sensibles à cette maladie.

3.2. Matériels et méthodes

3.2.1. Système d'étude

3.2.1.1. *Site expérimental*

La zone d'étude située sur la station expérimentale du Mas d'Adrien en Camargue est gérée par le CFR. Des essais techniques pour expérimenter de nouvelles pratiques culturales et conduire un programme de sélection variétale sont mis en place annuellement par le CFR. Ces essais sont également exploités par plusieurs instituts de recherche (CIRAD, INRA, CNRS, la Tour du Valat, etc.).

Dans le cadre du projet SEPYA, le CFR a mis en place, sur deux années consécutives, un essai expérimental sur deux parcelles voisines. La Figure 3.1 présente les photographies des parcelles expérimentales suivies lors des deux années de prélèvements.

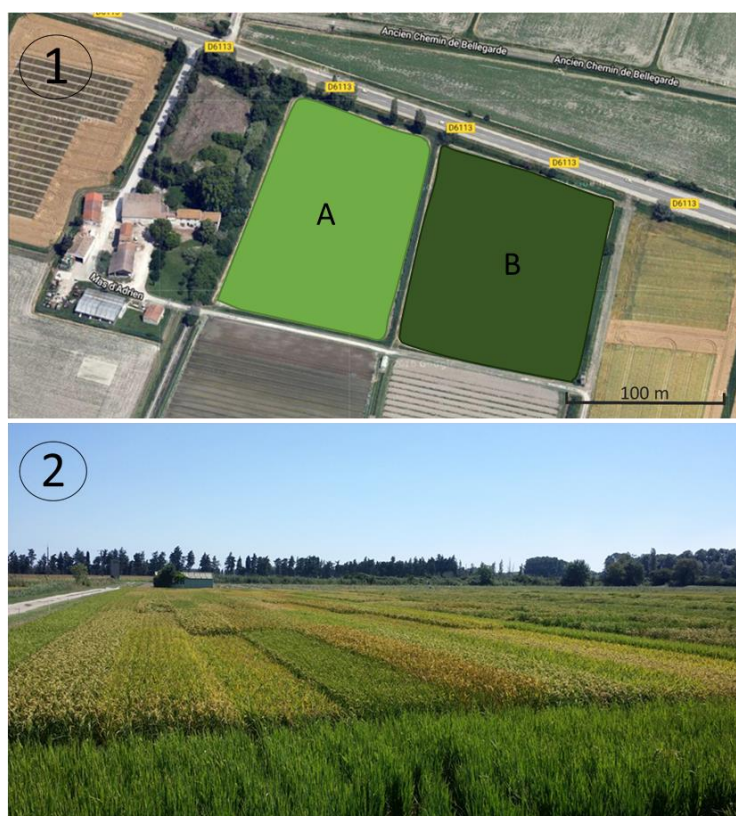


Figure 3.3 : 1) Station expérimentale du Mas d'Adrien : la parcelle « A » correspond au dispositif de l'année 2017 ; la parcelle « B » correspond au dispositif de l'année 2018 ; 2) Photo des microparcelles de la parcelle « B » au stade floraison en 2017

Les deux dispositifs mis en place respectivement en 2017 (parcelle A Figure 3.1 1) et en 2018 (parcelle B Figure 3.1 1), sont basés sur le même design expérimental. Au centre de chaque parcelle de 1.6 ha, ont été définis des blocs de microparcelles distribuées aléatoirement. Le terme « microparcelle » correspond à une surface rectangulaire de 9m² (6 mètres de long par 1,5 mètre de large) où a été semée une unique variété de riz, à une densité de 500 graines par m². Le semis des variétés de riz pour chaque microparcelle a été réalisé avec un semoir expérimental à huit socs. L'utilisation du semoir expérimental, lorsque la parcelle était asséchée, permet un semis de précision de huit rangs de riz espacés de 18 centimètres et évite les contaminations de semis entre les microparcelles. Les grains de riz ont tous été traités avec un agent mouillant, pour éviter la dérive des grains lors de la remise en eau de la parcelle après le semis. La mise en eau de la parcelle, a été réalisée par irrigation bihebdomadaire avec l'eau du Rhône. La parcelle a été maintenue en eau jusqu'à la fin du cycle de production du riz. Le dispositif expérimental a été divisé en sous parties, distribuées de façon homogène en 2017 ou de façon aléatoire en 2018, composées de rangs de microparcelles et chacune de ces sous parties a reçu un traitement différent (Figure 3.2). Trois niveaux de fertilisation azotée des sols ont été testés : un niveau normal de 150 unités d'azote (niveau recommandé aux agriculteurs en Camargue, parcelles en vert dans la Figure 3.2) ; un niveau faible avec 50 unités de moins que le niveau recommandé (parcelles en bleu, Figure 3.2) et un niveau élevé avec 50 unités de plus que le niveau recommandé en Camargue (parcelles en jaune, Figure 3.2). Par ailleurs, la moitié du dispositif a été exposé à des plants infectés par *Pyricularia oryzae*. Le dispositif global est répété deux fois sur la même parcelle.

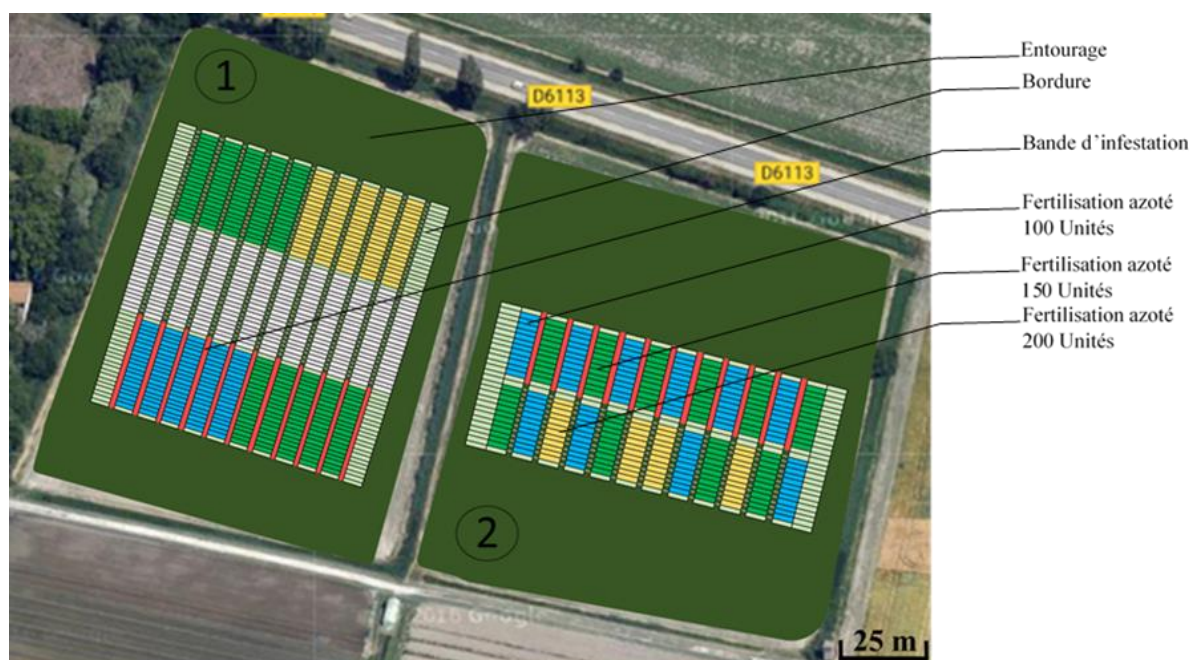


Figure 3.4 : Schéma des dispositifs expérimentaux de 2017 (1) et 2018 (2) : les dispositifs sont constitués d'un quadrillage de blocs de microparcelles entourées par une bordure semée avec la variété de riz (Manobi). Le reste des deux parcelles d'essais est semé avec une autre variété de riz (Gageron) et constitue « l'entourage ». Les couleurs jaunes, bleues et vertes des microparcelles correspondent aux différents niveaux de fertilisation azotée des sols et la couleur rouge entre les microparcelles correspond à la localisation des zones de repiquages des plants de riz infectés par *Pyricularia oryzae*.

En 2017 et 2018 les essais SEPYA avaient pour objectifs de tester deux variables sur le panel de variétés de riz semées : i) la fertilisation azotée initiale apportée au sol, ii) l'inoculation par *Pyricularia oryzae*. Pour des questions de moyen et de temps nous n'avons pas suivi l'effet de la fertilisation azotée sur la structuration du microbiote du riz. Nous avons choisi de centrer notre attention sur le niveau de fertilisation recommandé aux agriculteurs, qui est de 150 unités d'azote, et d'étudier dans ces conditions, les variations de structuration du microbiote endophyte au cours du cycle de développement du riz, entre différents génotypes de riz et sous pression de *Pyricularia oryzae*.

3.2.1.2. Espèce modèle : *Oryza sativa japonica* tempéré

L'essai SEPYA s'est focalisé sur la sous-espèce *japonica* tempéré d'*Oryza sativa* qui est traditionnellement cultivée en Camargue. Cinquante variétés de riz représentatives de la diversité française des riz *japonica* issues de programmes de sélection mis en place par le CFR, ont été choisies en 2017 pour leurs variabilités de sensibilité à la pyriculariose et leurs profils d'efficacité d'utilisation de l'azote. En 2018, 12 variétés parmi les 50 testées en 2017 ont été

retenues pour une deuxième année d'essai pour leur capacité à utiliser l'azote. Toutes les variétés semées ont un cycle court (entre 87 et 104 jours) et sont adaptées aux conditions climatiques de la zone d'étude. Ces variétés ont déjà fait l'objet d'analyses de génotypage par GBS et ont montré une diversité génétique faible au sein du groupe *japonica* tempéré d'Europe.

3.2.2. Echantillonnage

Nous avons sélectionné, parmi le panel de variétés des essais expérimentaux conduits en 2017 et 2018, six variétés présentant des niveaux contrastés de résistance à la pyriculariose en conditions de laboratoire (données UMR BGPI à partir de six souches de *Pyricularia oryzae*) (Tableau 3.1). Nous avons échantillonné en 2017 et 2018, ces six variétés sur chacune des parties du dispositif à plusieurs stades de développement. Deux microparcelles ont été échantillonnées par variété, avec trois plantes collectées en 2017, et deux plantes en 2018 (Figure 3.3).

Tableau 3.2 : Données disponibles sur l'estimation de la résistance à la pyriculariose en conditions de laboratoire

Variété	Numéro variété	Nombre de souches testées	Sensibilité sur le nombre de souches testées
GAGERON	SEPYA_006	6	0/6
MAMBO	SEPYA_008	6	2/6
PACO	SEPYA_010	6	3/6
RV000X	SEPYA_016	6	2/6
BRIO	SEPYA_017	5	2/5
RL000X	SEPYA_029	6	2/6

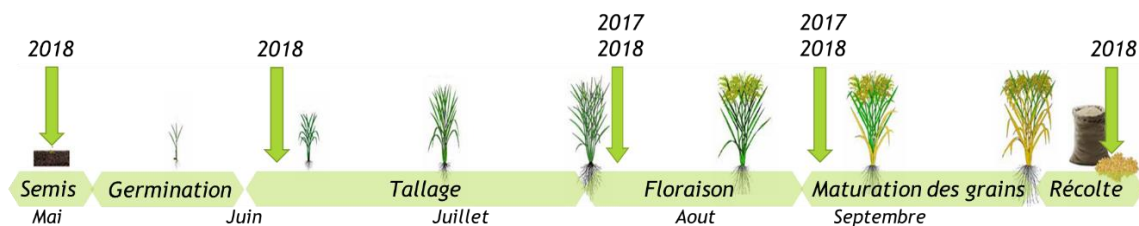


Figure 3.3 : Périodes d'échantillonnage sur les deux années de prélèvements

La caractérisation des communautés endophytes du riz pour les prélèvements de 2017, a été réalisée à deux stades de développement : au début de la floraison à 70 jours après semis (avant la mise en place des plants de riz inoculés par *Pyricularia oryzae* dans les bandes d'infestation) et au début de la maturation des grains à 110 jours après semis (1 mois après la mise en place des plants de riz inoculés par *Pyricularia oryzae* dans les bandes d'infestation sur la moitié du dispositif expérimental).

Le plan d'échantillonnage pour le réplica de notre étude en 2018, a évolué suite à la publication des travaux d'Edwards et de ses collaborateurs (Edwards et al., 2018) sur l'étude de la dynamique de structuration des communautés bactériennes, associées aux racines de riz. Ces travaux ont montré un changement de microbiote du rhizoplan et de l'endosphère des racines au cours des 7 à 8 premières semaines après la germination. Il se stabilise par la suite pour le reste du cycle de croissance du riz. En plus des prélèvements à 70 et à 110 jours après semis, nous avons choisi de rajouter dans le plan d'échantillonnage en 2018, un point de prélèvement supplémentaire à 47 jours après semis (soit environ 5 à 6 semaines après germination) correspondant à un stade précoce de tallage. Nous avons également décidé en 2018, d'analyser le microbiote des semences utilisées pour l'essai, ainsi que les grains récoltés pour chacune des six variétés suivies, afin d'estimer la part de la transmission verticale par semences dans la structuration du microbiote endophyte.

3.2.3. Préparation des échantillonnages

Deux compartiments de la plante ont été étudiés pour la caractérisation du microbiote endophyte du riz : la partie racinaire et la base de la tige. Après leur acheminement au laboratoire de la station expérimentale du CFR, les plants de riz ont été abondamment rincés à l'eau distillée. Nous avons tout d'abord éliminé les tiges secondaires du talle, les feuilles et l'ochréa (gaine qui entoure la tige). Nous avons ensuite prélevé le premier entre-nœud (en partant de la racine en conservant les deux nœuds de part et d'autre de l'entre-nœud) et prélevé 5 racines principales dans toutes leurs longueurs (Figure 3.4).

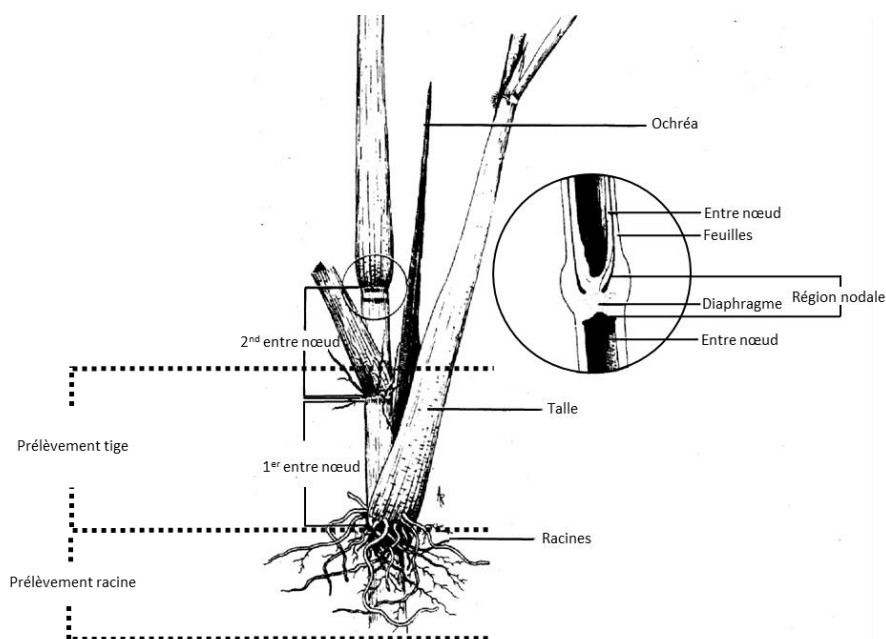


Figure 3.4 : Schéma des compartiments aérien et racinaire prélevés, schéma inspiré de (Bardenas, 1965)

En 2017, 144 échantillons de racines et de tiges ont été prélevés. En 2018, 216 échantillons de racines et de tiges ont été collectés. Enfin, 84 échantillons de graines ont été prélevés en 2018.

L'étude du microbiote endophyte du riz, nécessite d'éliminer les microorganismes présents à la surface des échantillons. Nous avons donc procédé immédiatement, après la préparation des échantillons, à une désinfection de leurs surfaces. Les échantillons prélevés ont été placés dans 20 ml d'une solution d'éthanol à 70% pendant 3 minutes sous agitation vortex à 2700 rpm (en plus de l'action chimique de l'éthanol, l'agitation permet un décrochage mécanique des microorganismes épiphytes). Les échantillons ont ensuite été placés 5 minutes dans 20 ml d'une solution de javel (à 2,5 degrés chloriques pour les échantillons de racines et 0.9 degrés chloriques pour les échantillons de tiges et graines), sous agitation vortex à 2700 rpm. Enfin une dernière étape de désinfection a été réalisée dans 20 ml d'une solution d'éthanol à 70%, pendant 30 secondes sous agitation vortex à 2700 rpm. Les échantillons ont ensuite été rincés 5 fois avec de l'eau distillée stérile, pendant 5 secondes, sous agitation vortex à 2700 rpm.

Sous hotte à flux laminaire, nous avons alors prélevé environ 0.3 g de tissu végétal. Pour les échantillons des tiges, nous avons prélevé la partie centrale entre les deux entre-nœuds. Les échantillons ont été prédécoupés en morceaux d'environ 1mm² avec un scalpel stérile, afin de faciliter les étapes de broyage. Les échantillons de tiges et de racines ont ensuite été placés dans

des tubes de 2 ml stériles, contenant deux billes métalliques de 2 mm et une bille céramique de 50 mm, avant de réaliser l'extraction de leurs ADNs génomiques totaux.

Trois feuilles de riz ont été par ailleurs prélevées sur chaque plant échantillonné, pour réaliser l'analyse de la diversité virale. Tous les échantillons ont été conservés individuellement à -80°C, en attendant les étapes d'extraction des acides nucléiques.

3.2.4. Analyse du virome par métagénomique VANA

L'analyse du virome du riz a été réalisée à partir des 144 et 216 échantillons de feuilles de riz respectivement collectées en 2017 et en 2018. Les protocoles de semi-purification VANA et de bio-informatique sont identiques à ceux décrits dans la partie 2.2.1 de ce manuscrit de thèse. Comme dans l'analyse de la partie 2.2.1, le premier seuil de détection a été établi en triplant le nombre maximal de reads viraux retrouvés dans les cinq témoins négatifs, chacun constitué d'un bout d'une feuille de canne à sucre, issue de graine indemne de virus (Fuzz). Ce premier seuil a ainsi été fixé à 318 reads (3 fois 106 reads retrouvés dans un des cinq témoins de canne à sucre). Le second seuil a été basé sur la présence d'au moins un contig phytoviral supérieur à 1000 pb pour considérer l'infection d'une plante par un phytovirus.

3.2.5. Analyse des communautés bactériennes par métagénomique ciblée

3.2.5.1. Extraction de l'ADN génomique total

Il existe de nombreux kits commerciaux permettant d'extraire l'ADN génomique d'échantillons de tissus végétaux. Dans le chapitre 2 de ce manuscrit de thèse, nous avons utilisé le kit d'extraction NucleoMag® plant de Macherey Nagel. Or, celui-ci s'est avéré être fortement contaminé par des ADNs d'origine bactérienne. La présence de contaminants bactériens, même à faible concentration, dans les réactifs d'extraction peut être très gênante dans le cas d'analyses d'échantillons à faible biomasse bactérienne. En effet, elle crée une compétition avec l'ADN bactérien des échantillons lors de l'étape d'amplification et réduit grandement la profondeur de séquences (Eisenhofer et al., 2019). La biomasse bactérienne attendue pour la caractérisation du microbiote endophyte du riz étant plus faible que celle observée dans le cadre de l'analyse des communautés microbiennes globales réalisée dans le chapitre 2, nous avons décidé d'utiliser un autre kit d'extraction d'ADN. Nous avons choisi le kit PowerMag® Microbial DNA Isolation Kit, couramment utilisé dans de nombreuses études de caractérisation du

microbiote et en particulier pour le microbiote endophyte (Mareque et al., 2018, Jha et al., 2019).

Pour l'extraction d'ADN, 0,25 g de tissu végétal pour les échantillons de tiges et racines ou 3 graines pour les échantillons de semences préalablement conditionnés avec les billes de broyages, ont été maintenus congelés en azote liquide et ont été broyés à froid avec un broyeur RETSCH MM 400 à la vitesse de 1800 oscillations par minute, pendant deux fois 30 secondes. Suite à cette étape, le broyat congelé a été directement utilisé pour l'extraction d'ADN.

L'extraction de l'ADN a été effectuée en accord avec le protocole d'utilisation du kit PowerMag® Microbial DNA Isolation Kit sur le robot d'extraction KingFisher™ Flex Purification System. Pour chaque plaque d'extraction, un témoin négatif a été ajouté (utilisation unique des réactifs du Kit). L'efficacité et la qualité de l'extraction ont été contrôlées par une migration des acides nucléiques extraits par électrophorèse sur gel d'agarose 1% dans du tampon Tris-Acetate-EDTA (TAE) 1X.

Les analyses de diversité des microorganismes des plantes peuvent être affectées par la co-amplification des séquences d'ADN chloroplastique et mitochondrial. Une forte homologie de séquence est présente entre les séquences du gène de l'ARNr 16S des bactéries et l'ADN mitochondrial et chloroplastique de la plante hôte. Cette homologie de séquence entraîne une compétition pour l'hybridation des amorces lors de l'étape d'amplification PCR. L'amplification par PCR avec des amorces universelles ciblant l'ARNr 16S des bactéries est donc inefficace lorsque l'ADN de la plante hôte est dominant par rapport à l'ADN bactérien. Dans ce travail sur les communautés endophytes, le ratio ADN hôte/ADN bactérien est très déséquilibré. Nous avons donc choisi d'utiliser des séquences d'acides nucléiques peptidiques (PNA) spécifiques pour réduire l'amplification d'ADN chloroplastique et mitochondrial (Fitzpatrick et al., 2018, Lundberg et al., 2013). Le choix de la région ciblée (V3-V4 l'ARNr 16S) et des conditions d'amplifications PCR sont identiques à celles utilisées dans le chapitre 2 de ce manuscrit de thèse. Il en va de même pour l'analyse des séquences réalisée par la technologie MiSeq d'Illumina avec les réactifs « MiSeq Reagent Kit V2, 500 cycles Illumina, San Diego, CA ».

Un réplica technique de chaque échantillon a été réalisé à partir de l'étape d'amplification par PCR afin de réduire les biais de la méthode de séquençage par métagénomique ciblée, chaque échantillon a donc été amplifié par PCR deux fois puis séquencé de façon indépendante.

Pour chaque plaque PCR, un témoin négatif d'amplification a été réalisé avec l'eau du mix PCR afin d'identifier d'éventuelles contaminations bactériennes lors de l'étape d'amplification par PCR.

3.2.5.2. *Traitement bio-informatique des séquences*

Les méthodes de traitements des données brutes de séquençage permettant d'obtenir une table d'Unités Taxonomiques Opérationnelles (OTUs) sont identiques à celles présentées dans le chapitre 2. Les tables d'OTUs contenant l'abondance de chaque OTU pour chaque échantillon et son assignation taxonomique, ont ensuite subi un traitement bio-informatique différent de celui présenté dans le chapitre 2.

Nous avons tout d'abord, soustrait des échantillons de la table d'OTUs, l'abondance de toutes les séquences présentes dans les témoins négatifs d'extraction et de PCR afin d'éliminer les séquences considérées comme contaminantes.

L'utilisation des acides nucléiques peptidiques bloquants (PNA) permet de réduire l'amplification de l'hôte lors de la réaction PCR, mais leur action reste limitée lorsque le ratio ADN hôte/ADN bactérien est trop déséquilibré. Or, il s'est avéré que les communautés bactériennes endophytes de la partie aérienne étaient tellement peu abondantes dans ces tissus que l'utilisation des peptides bloquants n'était pas suffisante pour éviter l'amplification de la plante. Nous avons donc appliqué des filtres différents pour le traitement des données de séquençage provenant des échantillons de racines et des données de séquençage issues des échantillons de graines et de tiges.

Pour le nettoyage et la normalisation de la table d'OTUs des échantillons de racines, nous avons éliminé toutes les séquences associées à la plante, et regroupé les deux répliques techniques en conservant seulement les OTUs présentes dans les deux répliques techniques. A l'issue de ce traitement, les échantillons ayant moins de 5 000 séquences ont été éliminés du jeu de données, car ils ont été considérés comme ayant une couverture de séquençage insuffisante.

Un même traitement (élimination des séquences de plante puis élimination des échantillons avec moins de 5 000 séquences) est impossible pour les échantillons de tiges car la grande majorité des échantillons ne passerait pas le filtre de qualité du nombre minimum de séquences après l'élimination des séquences de l'hôte. Dans notre analyse bio-informatique, il était donc important de différencier les échantillons ayant peu de séquences bactériennes en raison d'une faible biomasse bactérienne dans les tissus analysés, des échantillons ayant peu de séquences

bactériennes en raison d'un problème technique lors des étapes d'amplification et de séquençage (Figure 3.5). Les échantillons à faible biomasse bactérienne ont la particularité d'avoir de nombreuses séquences associées à la plante hôte, alors que des échantillons avec une qualité de séquençage insuffisante ont globalement peu de séquences. Pour conserver les échantillons à faible biomasse, nous avons regroupé toutes les séquences des OTUs associées taxonomiquement au chloroplaste et à la mitochondrie en deux OTUs. Comme nous n'avons pas éliminé les séquences associées à l'hôte, la profondeur de séquençage minimale acceptable a été fixée à 10 000 séquences. Tous les échantillons inférieurs à ce seuil ont été considérés comme de qualité insuffisante et ont été supprimés du jeu de données. Suite à ce traitement de données, la majorité des échantillons a pu être conservée pour l'analyse des communautés bactériennes endophytes, mais peu de séquences (4,7%) ont été assignées à des OTUs bactériennes. Devant le peu de séquences bactériennes dans ces échantillons, nous avons préféré fusionner les deux répliques techniques au lieu de ne garder que les OTUs communes entre les deux répliques. En effet, cela permet d'augmenter la profondeur de séquençage, avec le risque assumé de conserver quelques OTUs chimériques (Brooks et al., 2015). Une normalisation des profondeurs de séquençage par raréfaction de la table d'OTUs a ensuite été réalisée pour homogénéiser l'effort de séquençage de tous les échantillons. La raréfaction a été réalisée sur l'échantillon à la plus faible profondeur de séquençage après élimination des échantillons de qualité insuffisante.

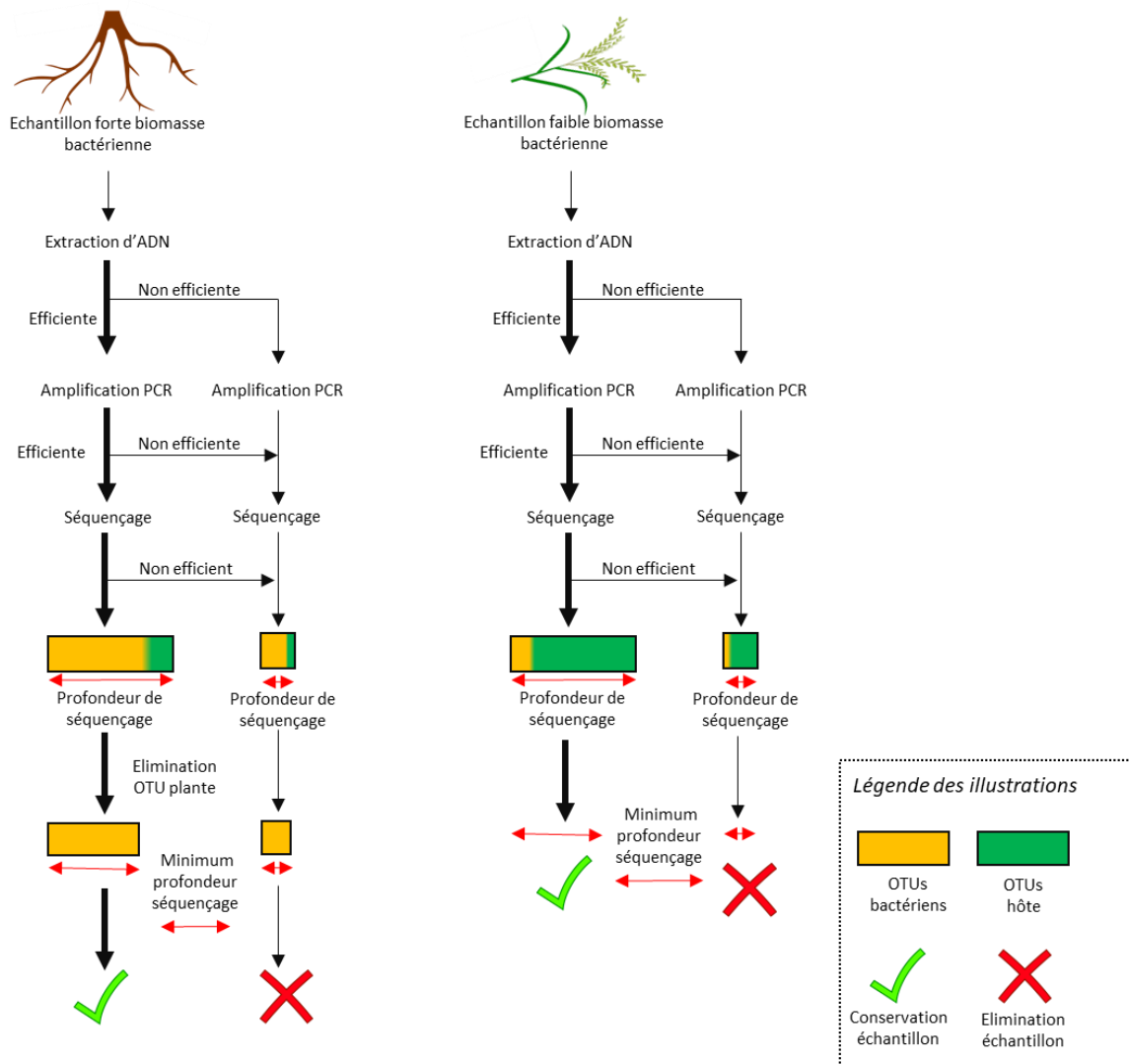


Figure 3.5 : Schéma explicatif pour justifier la différence de traitement bio-informatique entre les échantillons de racines et les échantillons de la partie aérienne

3.2.5.3. *Calcul des indices de diversité et détermination des facteurs de structuration des communautés bactériennes*

Plusieurs facteurs pouvant potentiellement façonner la structure des communautés bactériennes endophytes du riz ont été analysés : i) les compartiments de la plante, ii) les stades de développement de la plante, iii) l'année de prélèvement, iv) la variété de riz et v) la position dans la parcelle (effet bloc).

Les analyses de la diversité α et β nous ont permis d'estimer l'importance de chacun des facteurs suivis dans la structuration des communautés bactériennes endophytes. Les calculs des indices

de diversité α , richesse spécifique et indice de Shannon, ont été réalisés avec le Package R Phyloseq version 1.26.1 (McMurdie & Holmes, 2013) suivant la méthode présentée dans le chapitre 2 de ce manuscrit. L'analyse des variations de composition des communautés bactériennes entre les différentes conditions (diversité β), basée sur le calcul des distances de Jaccard et d'UniFrac non pondérés (Lozupone & Knight, 2005), a été réalisée avec le Package R Vegan version 5.2-3 (Oksanen, 2018). Elle est basée sur une analyse de la variance par permutation (PERMANOVA) et par une analyse en coordonnées principales (PCoA) selon les méthodes présentées dans le chapitre 2 de ce manuscrit. Nous avons choisi de travailler seulement avec les indices de diversité en ne prenant pas en compte l'abondance des séquences. Ces indices ne seront pas impactés par notre choix de conserver les OTUs associées à la plante hôte qui sont partagées par tous les échantillons.

Enfin, comme présenté dans le chapitre 2, l'identification des taxa significativement différenciellement abondants a été réalisée avec le package R Metacoder (Foster et al., 2017). Enfin, l'identification des taxa partagés chez une majorité des échantillons (avec plus de 80% de prévalence) appelés microbiote cœur a été réalisée avec le package R Microbiome.

Tous les codes R développés dans ce travail sont accessibles sur le site Guithub (https://github.com/P-alonso/micriobiota_analyses_SEPYA_camargue).

3.3. Résultats et discussions

L'objectif de notre étude était de comprendre la contribution de différents facteurs impliqués dans l'assemblage du microbiote endophyte du riz. Un facteur biotique, la présence du champignon phytopathogène *Pyricularia oryzae* pouvant contribuer aux variations de structure et de composition de ce microbiote, avait été intégré dans le design expérimental. En dépit de nombreuses tentatives d'inoculation de plusieurs souches de *Pyricularia oryzae* pendant les deux années de notre étude, aucun symptôme de maladie n'a pu être observé sur les variétés de riz sélectionnées dans l'essai. Nous n'avons donc pas étudié dans ce travail, l'effet de la présence de *Pyricularia oryzae* sur la structure des communautés bactériennes endophytes du riz et *vice versa*. Les répétitions avec ou sans *Pyricularia oryzae* n'étant pas différenciées par la présence du champignon, nous avons donc combiné les répétitions pour obtenir 4 microparcelles répétées par génotype.

3.3.1. Analyse du virome par métagénomique VANA

Comme pour la partie 2.1.1 de ce manuscrit de thèse, nous avons choisi deux seuils de détection pour éliminer les résultats faussement positifs liés au phénomène d'« index hopping » généré par la technologie Illumina (Valk et al., 2019). Nous avons choisi un seuil en triplant le nombre maximal de reads viraux retrouvés dans les témoins négatifs soit 318 reads viraux. Le rendement en reads viraux du run « Camargue » ayant été très supérieur à celui du run « Chine », a entraîné de nombreuses contaminations liées au phénomène d'« index hopping ». Le second seuil « contig » a été défini à 1 000 pb (alors que nous l'avions défini à 500 pb dans le chapitre 2) afin de s'assurer de l'élimination des faux positifs créés par le phénomène d'index hopping. Sur la base de ces deux seuils de détection, notre étude de métagénomique virale a permis d'identifier un seul virus de plante appartenant à la famille *Endornaviridae* présent dans plusieurs échantillons. Les différents contigs de ce virus, assemblés à partir des plantes contenant des reads d'*Endornaviridae*, ont tous présenté des identités nucléotidiques et protéiques élevées (>90%) avec *Oryza sativa alphaendornavirus*.

Bien que les analyses réalisées avec les deux seuils de détection donnent des résultats légèrement différents, les conclusions obtenues à partir des deux traitements sont similaires. La prévalence de ce virus est variable en fonction des variétés de riz et des stades de développement, mais les niveaux de prévalence virale sont similaires pour les deux années de prélèvement (Tableau 3.2 et 3.3 et Figure 3.6). D'après un test exact de Fisher avec une p-value inférieur à 0,05, il y a de façon significative, plus de plants infectés par *Oryza sativa alphaendornavirus* (OsEV) dans les variétés de riz Brio et RL000X aux trois stades phénologiques étudiés (tallage, floraison et maturation) que dans les 4 autres variétés testées. Les variétés Brio et RL000X sont deux variétés d'*Oryza sativa japonica* tempérées. La variété Brio est une variété commerciale italienne et la variété RL000X est une variété en cours de sélection par le Centre Français du Riz dont le parent femelle est la variété italienne Centauro. Une analyse de la diversité génétique de variétés à partir de 391 accessions de riz mondiaux comprenant les variétés Centauro et Brio (Biscarini et al., 2016), montrent un lien de parenté entre Centauro et Brio. La transmission d'OsEV est exclusivement verticale avec un taux de transmission très élevé pour les espèces *Oryza sativa japonica* (93% de taux de transmission avec des ovules infectés et 89% de taux de transmission avec des pollens infectés) (Horiuchi et al., 2003). Nous pouvons supposer que la présence du virus en Camargue pourrait être liée à des programmes de sélection de variétés sur la base de graines importées infectées par OsEV (Horiuchi et al., 2003, Valverde et al., 2011). La détection systématique d'OsEV dans plus de

50% des échantillons des variétés Brio et RL000X aux trois stades de développement, conforte par ailleurs les résultats de l'équipe de Moriyama qui a rapporté une charge virale d'OsEV constante dans tous les organes de la plante et à tous les stades de développement (Fukuhara et al., 2006). Par ailleurs, bien que nous n'ayons pas réalisé de relevés systématiques de symptômes, aucun symptôme n'a été identifié visuellement sur les 360 plants prélevés dans notre étude.

Tableau 3.2 : Analyse de la dynamique de prévalence virale l'oryza sativa alpha endornavirus pour les stades de développement floraison et maturation des grains en 2017 et tallage, floraison et maturation des grains en 2018 pour les six variétés de riz étudiées. Résultats avec le seuil de positivité fixé à au moins trois fois le nombre maximum de reads retrouvé dans un des témoins négatifs soit 318 reads minimum. En gras est représenté les prévalences virales de plus de 50%

Stade développement		Brio	Gageron	Mambo	Paco	RL000X	RV000X
2017	Floraison	91,7% (11/12)	33,3% (4/12)	16,7% (2/12)	8,3% (1/12)	83,3% (10/12)	16,7% (2/12)
	Maturation des grains	100% (12/12)	0% (0/12)	8,3% (1/12)	0% (0/12)	83,3% (10/12)	8,3% (1/12)
2018	Tallage	58,3% (7/12)	0% (0/12)	8,3% (1/12)	16,7% (2/12)	50,0% (6/12)	0% (0/12)
	Floraison	91,7% (11/ 12)	25,0% (3/12)	0% (0/12)	0% (0/12)	91,7% (11/12)	0% (0/12)
	Maturation des grains	91,7% (11/ 12)	0% (0/12)	0% (0/12)	8,3% (1/12)	100% (12/12)	0% (0/12)

Tableau 3.3 : Analyse de la dynamique de prévalence virale d'*oryza sativa* alphaendornavirus pour les stades de développement floraison et maturation des grains en 2017 et tallage, floraison et maturation des grains en 2018 pour les six variétés de riz étudiées. Résultats avec le seuil de positivité d'une séquence virale basé sur la présence d'au moins un contig phytoviral supérieur à 1000 paires de bases. En gras sont représentées les prévalences virales de plus de 50%

Stade développement	Brio	Gageron	Mambo	Paco	RL000X	RV000X
2017 Floraison	91,7% (11/12)	33,3% (4/12)	16,7% (2/12)	8,3% (1/12)	83,3% (10/12)	25,0% (3/12)
Maturation des grains	100% (12/12)	0% (0/12)	8,3% (1/12)	0% (0/12)	83,3% (10/12)	8,3% (1/12)
2018 Tallage	75,0% (9/12)	0% (0/12)	16,7% (2/12)	16,7% (2/12)	100% (12/12)	0% (0/12)
Floraison	91,7% (11/12)	25,0% (3/12)	0% (0/12)	0% (0/12)	91,7% (11/12)	8,3% (1/12)
Maturation des grains	91,7% (11/12)	0% (0/12)	8,3% (1/12)	8,3% (1/12)	100% (12/12)	0% (0/12)

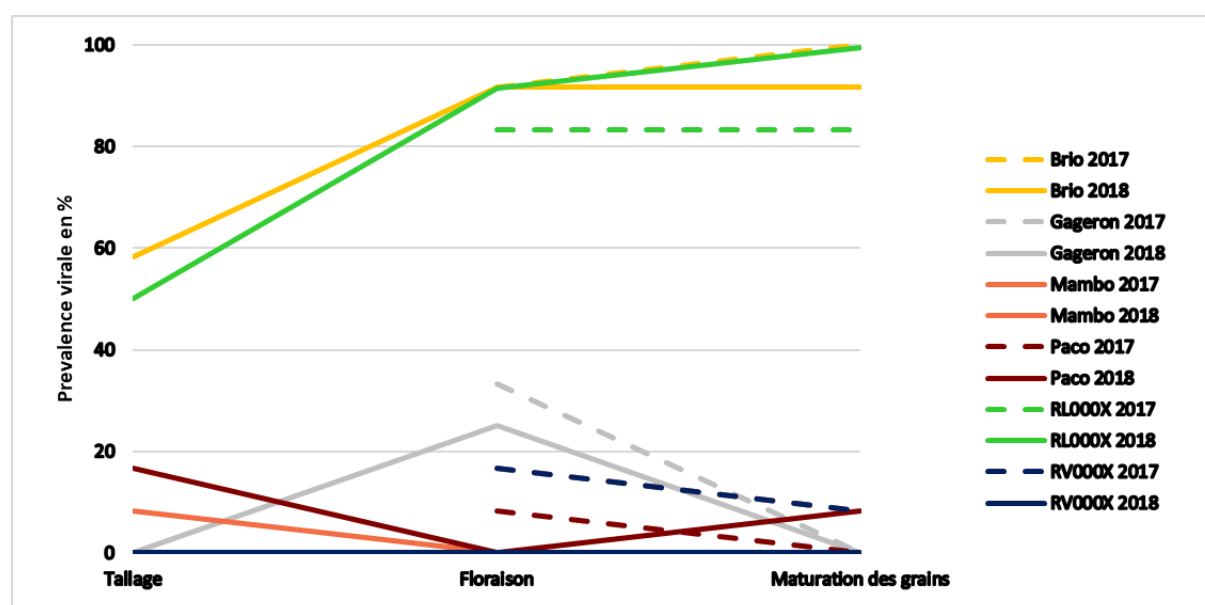


Figure 3.6 : Représentation de la dynamique de prévalence virale d'*oryza sativa* alpha endornavirus, l'année de prélèvement 2017 est représentée en pointillés et l'année de prélèvement 2018 en trait plein. Chaque couleur représente les différentes variétés de riz étudiées. Résultats avec le seuil de positivité fixé à au moins trois fois le nombre maximum de reads retrouvé dans un des témoins négatifs soit 318 reads minimum.

Les phytovirus de la famille *Endornaviridae* sont considérés comme des virus persistants, transmis verticalement, et n'ont généralement pas d'influence connue sur le phénotype de leur plante hôte (Fukuhara, 2019) . Une exception tout de même, le *vicia faba endornavirus*, est

connu pour produire une stérilité mâle des plants de fèves qu'il infecte (Mertens et al., 2005). Longtemps, les endornavirus ont été classés parmi les virus à double brin ARN (dsRNA). Le dernier rapport de l'ICTV les a reclassés parmi les virus à simple brin ARN (ssRNA) sur la base de données phylogénétiques, considérant la forme double-brin comme étant un intermédiaire répliatif (Fukuhara, 2019). *Oryza sativa* alphaendornavirus est un endornavirus appartenant au genre *Alphaendornavirus*. Découvert au Japon en 1999 (Moriyama et al., 1999b) il a été décrit comme étant présent chez de nombreuses variétés de riz au Japon sans présenter de symptômes visibles sur les plantes. Il se transmet exclusivement de façon verticale par le biais du pollen ou des graines (Fukuhara et al., 2006, Fukuhara, 2019). A notre connaissance, ce virus n'a été reporté qu'au Japon sur riz cultivé et sauvage (*O. rufipogon*) et sur quelques variétés aux Etats-Unis (Valverde et al., 2011). Il n'a jamais fait l'objet d'observations en Europe. Par ailleurs, OsEV n'a été observé que sur des variétés de riz de la sous espèce *japonica* et ne l'a jamais été à partir de variétés de la sous-espèce *indica* (Fukuhara, 2019).

OsEV est donc considéré comme un virus persistant qui se maintient à environ 100 copies par cellule tout au long du cycle du riz (Fukuhara, 2019). Ce nombre de copies virales assez faible est semble-t-il, le résultat d'une activation de la machinerie cellulaire du RNA silencing qui empêche le virus de se multiplier en grande quantité, mais « tolère » sa persistance à une faible concentration dans l'hôte (Fukuhara, 2019). Il a ainsi été proposé que cette régulation permet au virus de persister, puis de se transmettre verticalement sans affecter de façon majeure, la fitness de l'hôte (Fukuhara, 2019). Ces virus sont de fait, considérés comme étant symbiotiques de leurs hôtes et sont généralement différenciés des virus aiguë (acute virus en anglais) qui ont privilégié une stratégie de transmission horizontale et des activités intracellulaires de virulence pour s'accumuler dans l'hôte afin de maximiser leurs transmissions (Roossinck, 2015).

La découverte de ce virus dans nos travaux en Camargue et l'observation de sa prévalence notable au sein des plants de riz collectés nous permettra de tester en fin de ce chapitre si la présence d'OsEV est significativement corrélée à des différences de composition microbienne des plants infectés ou non par ce virus. Il sera par ailleurs, intéressant de voir s'il existe une différence d'effet sur le microbiote entre ce virus persistant et le fijivirus du riz décrit dans le chapitre précédent comme étant un virus aiguë causant des symptômes importants sur son hôte.

3.3.2. Analyse du microbiote endophyte des racines, des tiges et des graines de riz en Camargue

Suite au séquençage métabarcoding des communautés bactériennes endophytes du riz, 279 échantillons de racines, 280 échantillons de tiges et 84 échantillons de graines ont été analysés. Nous avons produit pour l'ensemble des échantillons de racines, tiges et graines respectivement 8 188 344, 22 411 511 et 5 769 950 séquences passant les filtres de qualité. Ces éléments ont abouti à identifier par l'algorithme d'agrégation de séquence Swarm respectivement 65 754, 3 600 et 172 OTUs. Le détail de la distribution des séquences entre les compartiments végétaux pour les deux années de prélèvements est présenté dans le tableau 3.4.

Tableau 3.4 : Tableau indiquant le nombre d'échantillons de chaque prélèvement conservé pour l'analyse de la diversité bactérienne et distribution du nombre de séquences et d'OTUs obtenues après nettoyage et raréfaction des données

Compartiment	Prélèvement	Nombre d'échantillons	Nombre de séquences	Nombre d'OTUs
Graines	semence 2018	12	127 680	27
Graines	grain 2018	72	766 080	108
Racines	Stade floraison 2017	68	307 156	2 322
Racines	Stade maturation des grains 2017	67	302 639	3 078
Racines	Stade tallage 2018	48	216 816	1 870
Racines	Stade floraison 2018	48	216 816	2 402
Racines	Stade maturation des grains 2018	48	216 816	2 662
Tiges	Stade floraison 2017	68	925 820	1 022
Tiges	Stade maturation des grains 2017	68	925 820	1 857
Tiges	Stade tallage 2018	48	653 520	444
Tiges	Stade floraison 2018	48	653 520	426
Tiges	Stade maturation des grains 2018	48	653 520	817

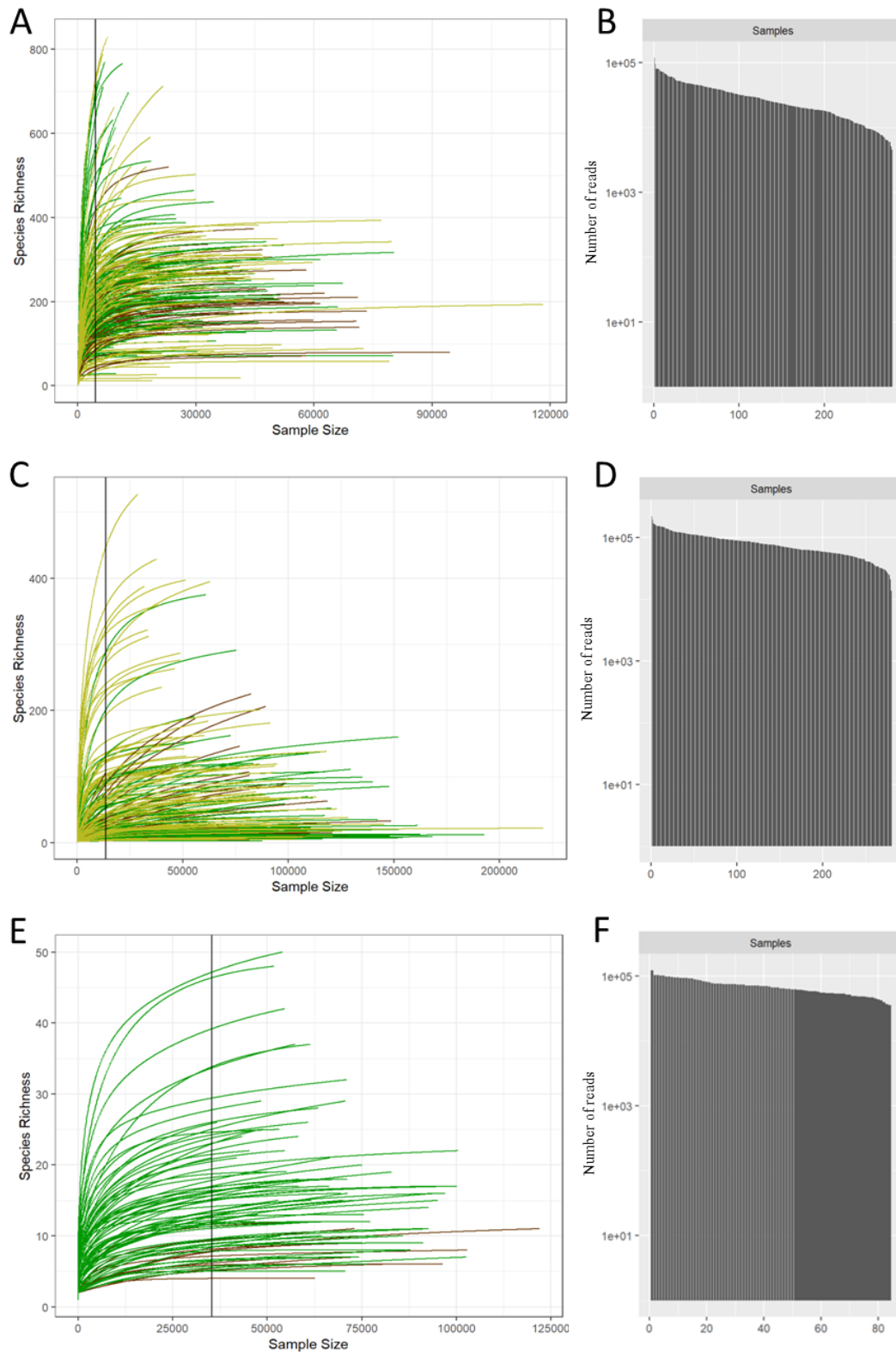


Figure 3.7 : Courbes de raréfaction des OTUs des communautés microbiennes endophytes et représentation des profondeurs de séquençages par échantillon selon les tissus végétaux. Les graphiques A et B correspondent aux échantillons de racines, C et D aux échantillons de tiges et E et F aux échantillons de graines. La droite verticale noire sur les graphiques A, C, et E correspond au seuil de normalisation pour les raréfactions des données.

Afin de savoir si nous avons produit un effort de séquençage suffisant pour couvrir la diversité des communautés bactériennes pour chaque tissu végétal analysé, nous avons tracé les courbes de raréfaction pour tous les échantillons analysés (Figure 2.7 A). L'effort de séquençage semble suffisant pour les échantillons de graines car toutes les courbes de raréfaction atteignent un plateau approximativement à partir de 25 000 séquences, indiquant qu'il y a peu de nouvelles OTUs attendues au-delà de ce seuil pour chaque échantillon. En revanche, la richesse bactérienne étant plus importante pour les échantillons de tiges et de racines, le plateau de la courbe de séquençage n'est pas atteint par certains échantillons. Cela suggère que notre analyse ne couvrira pas la totalité de la diversité bactérienne car certains taxa n'auront pas été séquencés, les taxa les plus rares seront mal échantillonnés.

La distribution des séquences entre les échantillons traduit un effort de séquençage relativement homogène (Figure 2.7 D/F). Malgré cette relative homogénéité nous avons procédé à un sous-échantillonnage aléatoire des séquences pour normaliser les efforts de séquençage entre les échantillons et pouvoir les comparer. Les seuils de raréfactions des données ont été fixés sur la plus petite profondeur de séquençage obtenue, pour chaque tissu après élimination des échantillons considérés comme de qualité de séquençage insuffisante (Figure 2.7 A/C/E). Les données ont été raréfiées après application des filtres de qualité des séquences et élimination des échantillons de faible profondeur soit 5 000 séquences pour l'analyse des communautés bactériennes sur la plante entière (racines, tiges, graines) et pour les communautés bactériennes du compartiment racinaire ; 10 000 séquences pour les communautés bactériennes de la tige de riz et 30 000 séquences pour les communautés bactériennes des grains de riz.

Après normalisation des profondeurs de séquençage par raréfaction, nous pouvons voir que les efforts de séquençage restent tous satisfaisants pour les échantillons de graines (Figure 2.7 E). La majorité de la diversité bactérienne après raréfaction des profondeurs de séquençages est également bien couverte pour les échantillons de tiges, à l'exception d'une dizaine d'échantillons ayant les richesses bactériennes les plus importantes (Figure 2.7 C). Les échantillons de racines, pour lesquels la profondeur de séquençage retenue par échantillon est la plus faible des trois compartiments végétaux, montrent une couverture partielle de la diversité bactérienne. (Figure 2.7 A). Ce résultat suggère que faute de profondeur de séquençage suffisante, une partie des OTUs les plus rares de la communauté bactérienne endophyte des racines ne sera pas prise en compte dans les analyses de la diversité bactérienne. Nous avons fait le choix pour notre analyse de ne pas augmenter le seuil du nombre de séquences avant la raréfaction pour ne pas éliminer trop d'échantillons et perdre en puissance statistique,

notamment dans les tests portant sur les différents indices de diversité qui seront calculés par la suite. Bien que ce choix de raréfaction ait entraîné une perte de sensibilité sur l'inventaire des communautés bactériennes pour les taxa les plus rares. La diversité des communautés bactériennes endophytes des racines restera bien estimée pour les taxa majoritaires (Weiss et al., 2017). Un réséquençage des échantillons les moins bien séquencés (peu de séquences et/ou forte richesse bactérienne), permettrait probablement d'améliorer les conclusions de ce chapitre en réalisant une analyse plus fine de la structure des communautés microbiennes.

3.3.2.1. Effet de différentes variables sur la structuration des communautés bactériennes endophytes du riz en Camargue

Sur deux parcelles expérimentales, nous avons suivi aux champs, l'effet du génotype du riz sur la structure des communautés bactériennes endophytes du riz à différents stades de développement de la plante. Le travail sur des parcelles expérimentales avec une randomisation du design expérimental, avait pour objectif d'essayer de contrôler au maximum l'homogénéité des variables environnementales sur l'essai, tout en restant en condition aux champs.

En plus de l'effet du génotype de riz sur la structure des communautés bactériennes endophytes du riz, nous avons suivi dans cette étude les variables suivantes : compartiment de la plante ; stade de développement de la plante ; année de culture et localisation de l'échantillon dans la parcelle (effet bloc). Par ailleurs, l'identification par l'analyse de la diversité virale de la présence élevée d'un phytovirus du genre *Alphaendornavirus* (39% et 23% de prévalence respectivement en 2017 et 2018) a permis de tester si cette variable pouvait avoir un effet sur la structure et la composition du microbiote des échantillons de riz.

Une première analyse réalisée sur l'ensemble du jeu de données, a permis d'estimer l'importance de chaque facteur dans la structuration du microbiote endophyte du riz. L'analyse de PERMANOVA sur l'indice de Jaccard, montre un effet significatif majeur des compartiments de la plante (tige, racine, graine) sur la structure des communautés bactériennes endophytes du riz. Le compartiment d'origine des échantillons explique 14% de la variance de la structure des communautés bactériennes observées. Ensuite, dans un ordre décroissant et de façon significative arrive : le stade de développement de la plante (2%) ; l'année de prélèvement (1%) puis la variété de riz (1%) (Tableau 3.5). Les variables « bloc » et « infection par l'endornavirus » n'ont pas montré d'effet significatif sur la structuration des communautés bactériennes endophytes du riz. Une forte part des variations (82%) reste non expliquée par les variables suivies (Tableau 3.5). Ces résultats obtenus dans des conditions environnementales

homogènes sont concordants avec des résultats déjà observés sur les communautés microbiennes globales (communautés endophytes et épiphytes confondues) de la canne à sucre (Hamonts et al., 2018) où le compartiment d'origine des échantillons expliquait une part importante des variations de structure des communautés bactériennes (19%), l'âge de la plante et son génotype n'expliquant chacun que moins de 2% des variations. Plusieurs études sur d'autres plantes (agaves, panic érigé) ont montré des résultats similaires avec à la fois, un effet majeur du compartiment de la plante et un effet faible du génotype sur la structuration du microbiote (Singer et al., 2019, Coleman-Derr et al., 2016). La forte part de variations de structuration des communautés bactériennes qui reste inexpliquée, est aussi cohérente avec les différentes études citées qui ont toutes un résiduel de variations des analyses de PERMANOVA toujours élevé. L'article de référence dans l'étude des facteurs de structuration du microbiote chez les graminées, montre ainsi un résiduel proche de celui que nous observons avec 72% de la variance de la structure des communautés microbiennes qui reste non expliquée (Hamonts et al., 2018). L'analyse du microbiote racinaire chez le riz séparé en trois compartiments (rhizosphère, rhizoplan et endosphère) a aussi permis de mettre en évidence la contribution relative de différents facteurs dans l'assemblage de ce microbiote (Edwards et al., 2015, Edwards et al., 2018). En conditions contrôlées, le compartiment racinaire (rhizosphère, rhizoplan et endosphère) ; le type de sol et le génotype du riz ont expliqué de façon décroissante les variations observées chez le microbiote du riz alors qu'aux champs les facteurs importants ont été le site géographique (séparés de plus 100 Kms), puis le compartiment racinaire et les pratiques culturales (conventionnel vs. biologique) (Edwards et al., 2015). Dans une étude postérieure de suivi du microbiote racinaire du riz aux champs, Edwards et coll. (2018) ont confirmé l'importance du compartiment racinaire dans les variations du microbiote, puis le site géographique (séparés de plus de 100 Kms), le stade de développement du riz, puis l'année et le génotype du riz qui ont aussi contribué de façon significative aux variations du microbiote du riz (Edwards et al., 2018). Par ailleurs, en comparant les compartiments souterrain et aérien, le microbiote de la rhizosphère du riz a été clairement séparé de celui de la phyllosphère en termes de composition (Knief et al., 2012). La forte part de la variation de structure des communautés microbiennes non expliquée dans toutes ces études, dont nos travaux sur le microbiote du riz en Chine et en Camargue, demande à être approfondie par une meilleure connaissance et maîtrise des variables responsables de la différence de structure des communautés microbiennes associées aux plantes. Plusieurs travaux ont souligné l'importance de phénomènes stochastiques dans l'explication des variations de structure du microbiote chez les animaux et des communautés microbiennes en général (Sieber et al., 2019, Nemergut et al.,

2013). A titre d'exemple, nous ne pouvons pas écarter qu'une partie de la structure des communautés microbiennes associées aux plantes ne soit pas la conséquence d'événements de colonisations aléatoires provenant de l'environnement la plante (Zhou & Ning, 2017, Maignien et al., 2014).

Tableau 3.5 : Estimation de l'influence des différents facteurs sur la structure des communautés bactériennes endophytes du riz à l'aide d'un test de PERMANOVA basé sur l'indice de Jaccard (10 000 permutations). Les résultats significatifs avec une P-value inférieure à 0,05 sont représentés en gras, la valeur R² désigne la proportion de la variance qui peut être expliquée par chacune des variables étudiées

	R ²	Pr (< F)
Plant compartment	0.14	<0.0001
Developement stage	0.02	<0.0001
Sampling year	0.01	<0.0001
Rice variety	0.01	0.0095
Bloc	0.00	0.0927
<i>Alpha Endornavirus</i>	0.00	0.2645
Residual	0.82	

3.3.2.2. *Les communautés bactériennes endophytes du riz sont différentes entre les organes de la plante*

- Diversité α : richesse bactérienne

Le compartiment de la plante est le facteur de structuration principal des communautés bactériennes endophytes du riz dans notre étude. La richesse des communautés endophytes bactériennes suit un gradient décroissant pour respectivement les compartiments racines, tiges et graines, (Figure 3.8).

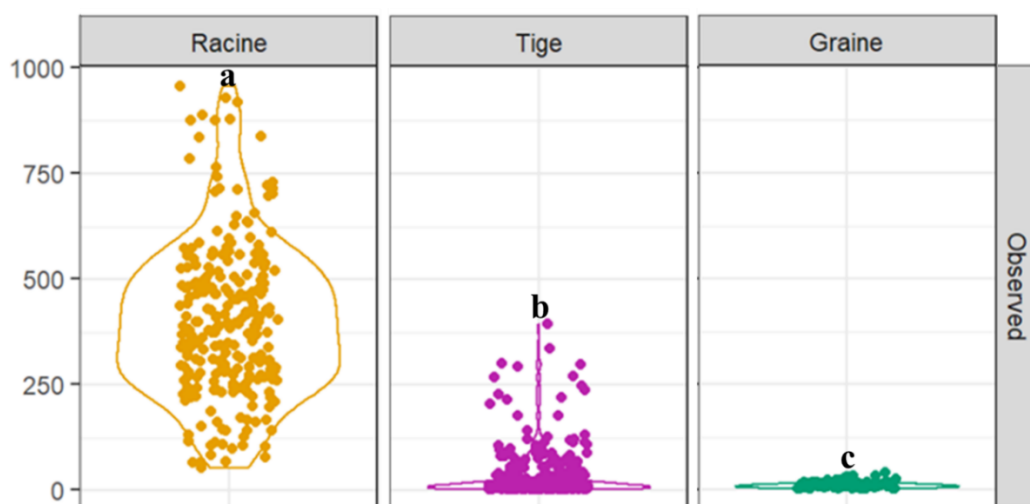


Figure 3.8 : Diagramme en violon de la richesse bactérienne endophyte dans les différents organes de la plante, les P-values inférieure à 0.05 du test de Tukey HSD tests sont représentées par les lettres a, b et c. La richesse microbienne est significativement différente dans chacun des compartiments de la plante avec un gradient de richesse décroissant de la racine à la graine.

Bien qu'à notre connaissance aucun travail récent n'a étudié les variations de structures des communautés endophytes du riz entre la partie souterraine et la partie aérienne, Mano et Morisaki ont reporté en 2008 (Mano & Morisaki, 2008) à partir de travaux réalisés par mise en culture, que la richesse bactérienne endophyte du riz était plus importantes dans les racines que dans les feuilles et les graines. Cette revue présente aussi les voies d'entrées des bactéries dans les tissus de la plante. Le sol y est présenté comme un réservoir de bactéries qui sont susceptibles de coloniser la surface de la racine et de pénétrer par la suite dans le compartiment endophyte de celle-ci. Certaines bactéries endophytes des racines peuvent être transportées vers les organes aériens de la plante via le système vasculaire de celle-ci. Plus récemment, des travaux de caractérisation du microbiote entre les compartiments de la plante réalisés sur des modèles non riz ont démontré par des approches de métagénomique que la richesse microbienne (bactéries, champignons et archées) était plus importante dans les racines que dans la phyllosphère (Ma et al., 2013, Qian et al., 2019, Park et al., 2017, Coleman-Derr et al., 2016). De plus, il a été montré que des communautés microbiennes du sol contribuaient au microbiote de la phyllosphère (Copeland et al., 2015, Bai et al., 2015). Ainsi, le modèle qui ressort de ces travaux propose qu'un ensemble de microorganismes du sol migrent de la rhizosphère à l'endosphère des racines, puis qu'une fraction de ces microorganismes, potentiellement sélectionnée par la plante, est alors transportée vers l'endosphère des parties aériennes via le xylème (Compant et al., 2010, Compant et al., 2019). En complément de cette voie « interne »

de transfert des bactéries, il a aussi été montré qu'une partie des microorganismes de la phyllosphère peut pénétrer dans la plante au niveau de ses parties aériennes et contribuer au microbiote endophyte (Vorholt, 2012).

Pour expliquer la plus forte richesse bactérienne endophyte associée aux racines de riz, nous ne pouvons pas écarter l'éventuelle désinfection incomplète des échantillons. L'augmentation de richesse bactérienne endophyte observée entre les tiges et les racines pourrait être expliquée, dans une certaine mesure, par les difficultés rencontrées pour éliminer de la surface des échantillons les bactéries du rhizoplan (plus riche) ou leurs ADN résiduels (Wang et al., 2016, Sessitsch et al., 2012, Lundberg et al., 2012).

- Diversité β : structures bactériennes des différents compartiments de la plante

En complément de l'analyse de PERMANOVA sur la distance de Jaccard qui montre un effet significatif du compartiment de la plante sur la structure des communautés bactériennes endophyte du riz, les résultats de PCoA permettent de distinguer assez clairement trois groupes de communautés bactériennes correspondant chacun à un des trois compartiments analysés (Figure 3.9). Nous pouvons également noter une plus forte hétérogénéité dans la structure des communautés microbiennes endophytes pour les échantillons de graines et de tiges, par rapport aux communautés bactériennes des racines. Les parties aériennes des plantes sont plus exposées à des fluctuations environnementales (température, humidité, radiation UV, variabilité d'accès aux nutriments...). Cela conduit à des variations de compositions des communautés microbiennes de la phyllosphère (Bulgarelli et al., 2013). Cette hétérogénéité de la source épiphyte phyllosphérique pour les communautés endophytes aériennes et/ou les effets stochastiques, liés à la migration pour coloniser ces compartiments, ont pu contribuer à la plus forte variabilité observée chez les communautés endophytes aériennes.

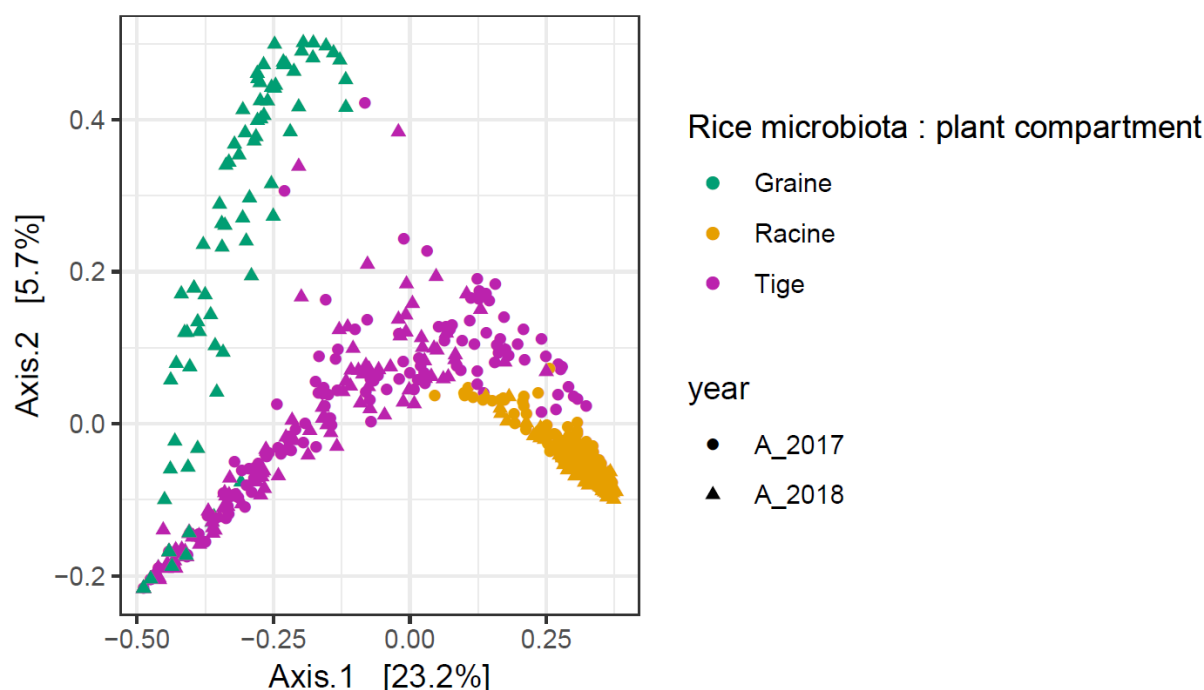


Figure 3.9 : Structure des communautés bactériennes endophytes du riz entre les compartiments végétaux, Analyse en composantes principales (PCoA) utilisant l'indice de Jaccard entre les échantillons. Les couleurs jaune, rose et verte correspondent respectivement aux échantillons de racines, tiges et graines. Les formes des points du graphique correspondent aux années de prélèvements.

- Abondance différentielle des taxa entre les compartiments de la plante

Nous avons ensuite réalisé des tests de rang pour identifier les différences d'abondances taxonomiques entre les communautés bactériennes de chaque compartiment végétal (Figure 3.10). A partir de l'arbre taxonomique global représentant les différentes OTUs de tous les échantillons, nous observons que le microbiote endophyte des racines renferme des taxa qui y sont significativement plus abondants que dans la tige ou dans la graine. Ces analyses montrent qu'aucun taxon bactérien n'est retrouvé de façon plus abondante dans les tiges par rapport aux racines. Les trois taxa qui apparaissent significativement plus abondants dans les tiges ne correspondent pas à des OTUs bactériennes, mais sont assignées à des séquences de l'hôte (Figure 3.10). L'exclusive surabondance de certains taxa dans les racines et l'absence de taxa significativement plus abondants dans les échantillons de tiges suggèrent une possible migration partielle du microbiote endophyte racinaire vers la tige (Figure 3.9).

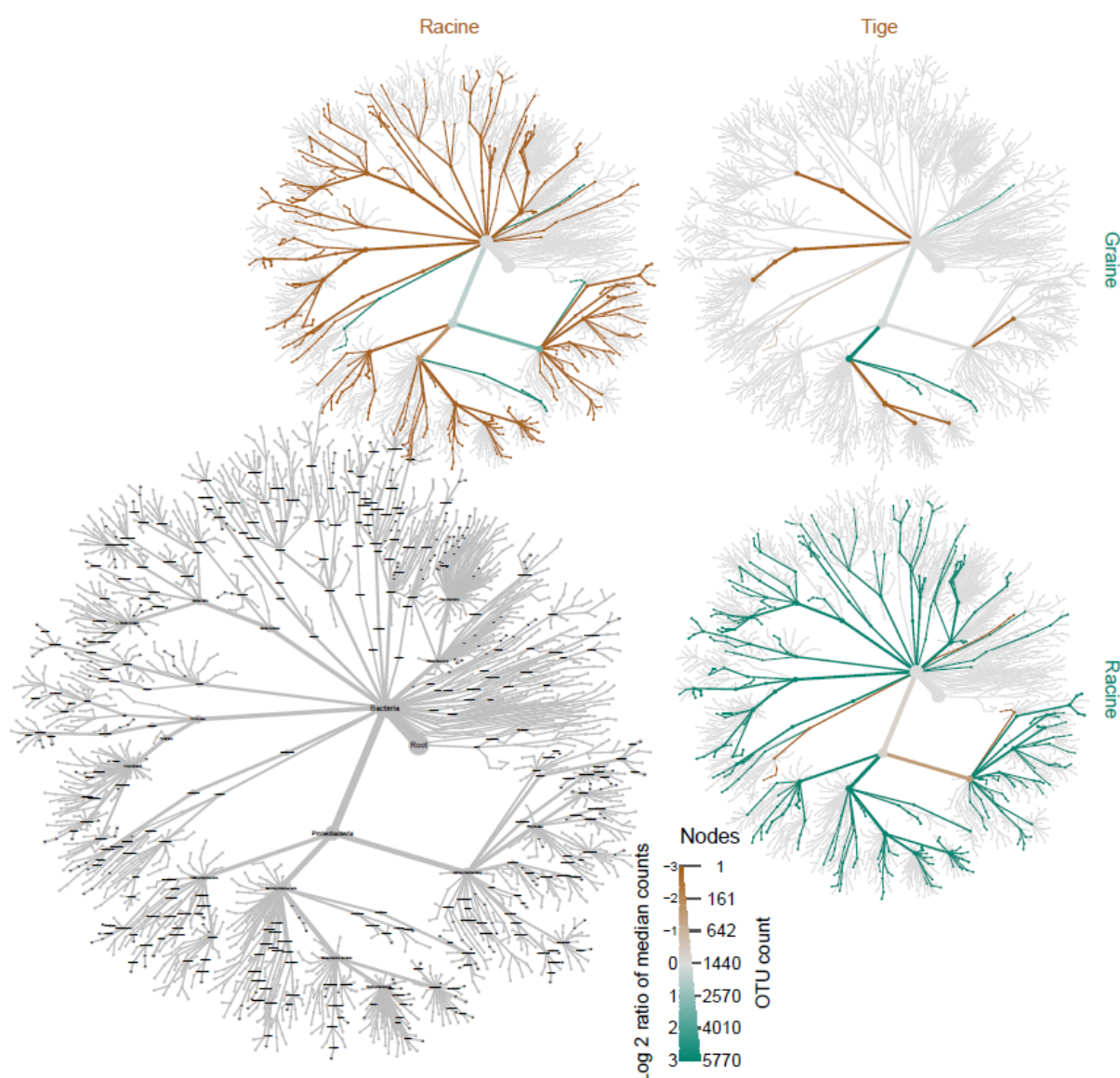


Figure 3.10 : Abondance différentielle des taxa entre les compartiments. L'arbre taxonomique gris représente la distribution taxonomique des bactéries endophytes du riz obtenu à partir de l'ensemble des échantillons analysés. La taille des nœuds correspond à l'abondance relative de chaque taxon. Cet arbre sert de clé pour identifier les taxa mis en évidence dans les « Heat tree différentiel » qui ne sont pas annotés. **Les arbres taxonomiques sont appelés « Heat tree différentiel » ils représentent les résultats des comparaisons par paires de l'abondance des taxa bactériens entre les compartiments végétaux.** La couleur de chaque taxon représente le rapport log-2 de la médiane des abondances observées avec chaque groupe (racines, tiges et graines) et correspond aux différentes comparaisons par paire réalisée. La couleur marron indique une abondance des taxa significativement (test de Wilcoxon, $p < 0,05$) plus élevée dans la variable positionnée en colonne lors de la comparaison par paire. La couleur verte indique une abondance des taxa significativement plus élevée dans la variable positionnée en ligne lors de la comparaison par paire. Tous les taxa de couleur grise ne présentent pas d'abondance relative significativement différente entre groupes comparés.

Ces analyses montrent par ailleurs, que trois OTUs bactériennes sont significativement plus abondantes dans les graines par rapport aux autres compartiments de la plante (Figure 3.10).

Ces trois OTUs sont environ 1 000 à 2 500 fois plus abondantes dans les graines que dans les échantillons de racines et environ 20 à 500 fois plus abondantes que dans les échantillons de tiges (toutes les données d'abondances relatives sont présentées en Annexe 1). Ces trois OTUs appartiennent à la classe des Gammaprotéobactéries. Dans la limite de notre assignation taxonomique, elles correspondent respectivement aux genres *Pantoea* et *Pseudomonas* ainsi qu'à un genre non identifié de la famille des *Enterobacteriaceae*. Toutes ont été décrites comme appartenant au microbiote cœur des graines de riz (Eyre et al., 2019, Zhang et al., 2019, Hardoim et al., 2012, Mano & Morisaki, 2008). La forte abondance de ces trois OTUs dans les échantillons de graines suggère la sélection de ces OTUs par la plante. Néanmoins, l'importance biologique de ces trois OTUs pour la plante reste difficile à estimer. Les genres de ces OTUs contiennent de nombreuses espèces bactériennes pouvant être soit pathogènes pour les plantes (Pili et al., 2016, Yan et al., 2010) soit bénéfiques en agissant comme agent de bio contrôle face à des microorganismes pathogènes (Feng et al., 2003, Wu et al., 2016) ou en favorisant la croissance des plantes (Zhang & Birch, 1997, Zhang et al., 2010b).

3.3.2.3. *La structuration des communautés bactériennes endophytes du riz varie avec le stade de développement du riz*

Nous avons analysé la dynamique de structuration des communautés bactériennes endophytes chez six variétés à différentes dates en 2017 (70 et 110 jours après semis) et en 2018 (47, 70 et 110 jours après semis). Bien que les différentes variétés n'aient pas toutes exactement le même cycle de développement entre elles et en fonction des saisons (Tableau 3.6), les prélèvements à 47 jours correspondaient pour toutes les variétés à un stade précoce de tallage ; les prélèvements à 70 jours au stade floraison et les prélèvements à 110 jours au stade maturation des grains.

Tableau 3.6 : Longueur du Cycle Semis-Epiaison (CSE) pour les six variétés suivies sur les deux années de prélèvements

Variétés	CSE 2017	CSE 2018
GAGERON	96,5 jours	104 jours
MAMBO	92,5 jours	100 jours
PACO	85,5 jours	98 jours
RV000X	90 jours	98 jours
BRIO	87 jours	99 jours
RL000X	87 jours	97 jours

- Dynamique de structuration des communautés bactériennes endophytes des racines

Des différences significatives de richesses bactérienne ont été observées entre les stades de développement du riz (Figure 3-11 C/D). Les variations de dynamiques de richesses sont différentes entre les compartiments végétaux. Ce résultat est confirmé pour les deux années de prélèvements. Les analyses de PERMANOVA (sur les indices de Jaccard et UniFrac non pondéré par l'abondance des OTUs) pour les échantillons de racines prélevés en 2017 montrent que le stade de développement (passage du stade floraison au stade de maturation des graines) contribue faiblement, mais de manière significative à l'explication des variations de structure des communautés bactériennes endophytes (3,2% des variations expliquées) (Tableau 3.7). L'analyse de PCoA réalisée sur l'indice d'UniFrac non pondérée illustre ces résultats en suggérant une légère différenciation des communautés bactériennes (Figure 3.12).

Tableau 3.7 : Estimation de l'influence du stade de développement du riz sur la structure des communautés bactériennes endophytes des racines et des tiges de riz par une analyse de PERMANOVA basée sur l'indice de Jaccard et l'indice UniFrac non pondéré (10 000 permutations). En gras sont représentés les résultats significatifs avec une P-value inférieure à 0,05, la valeur R² désigne la proportion de la variance qui peut être expliquée par chacune des variables étudiées

		Jaccard		UniFrac non pondéré	
		R ²	Pr (< F)	R ²	Pr (< F)
Racine 2017	Stade de développement	0.032	1e-04	0.028	1e-04
Racine 2018	Stade de développement	0.090	1e-04	0.095	1e-04
Tige 2017	Stade de développement	0.072	1e-04	0.087	1e-04
Tige 2018	Stade de développement	0.066	1e-04	0.097	1e-04

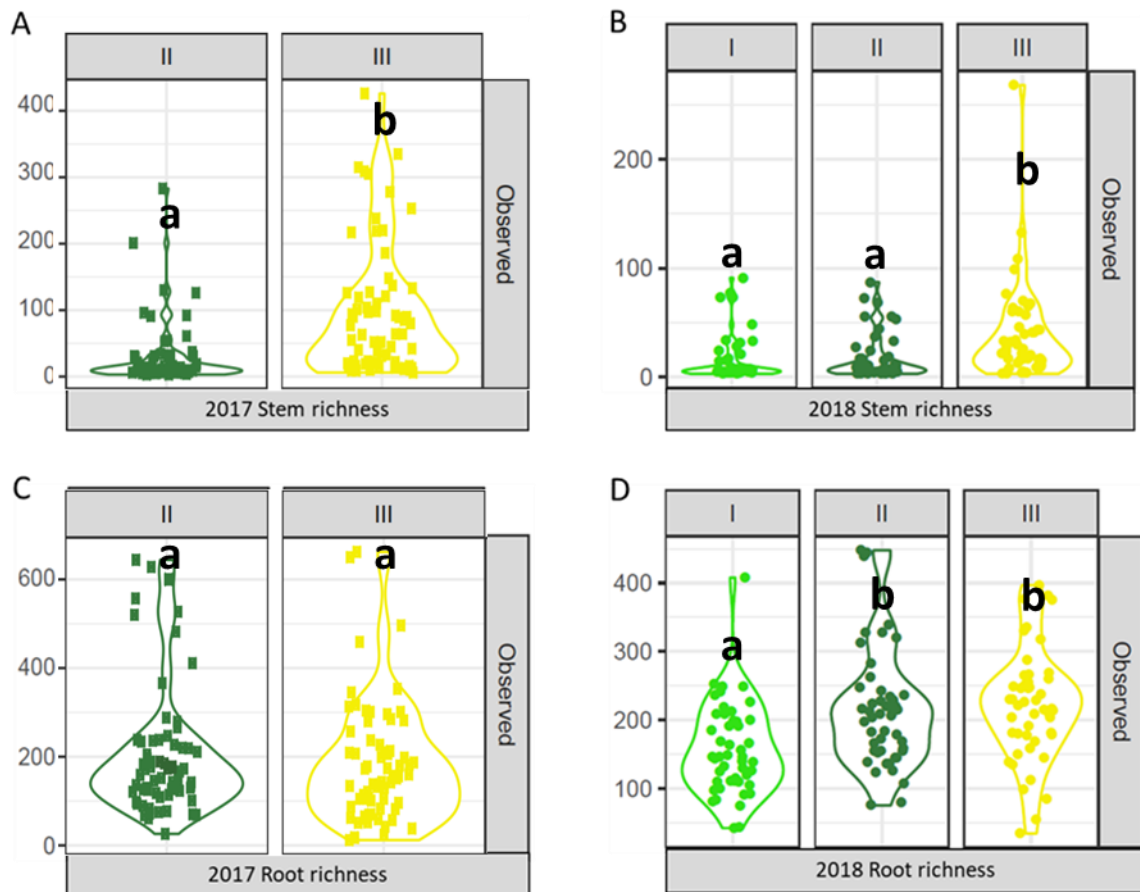


Figure 3.11 : Diagramme en violon de la richesse bactérienne endophyte des tiges (A ; B) et des racines (C ; D) entre les stades de développement de la plante pour les années de prélèvement 2017(A ; C) et 2018 (B ; D), Les diagrammes (I) correspondent au stade de développement du riz tallage, (II) au stade de développement floraison et (III) au stade de développement maturation des grains.

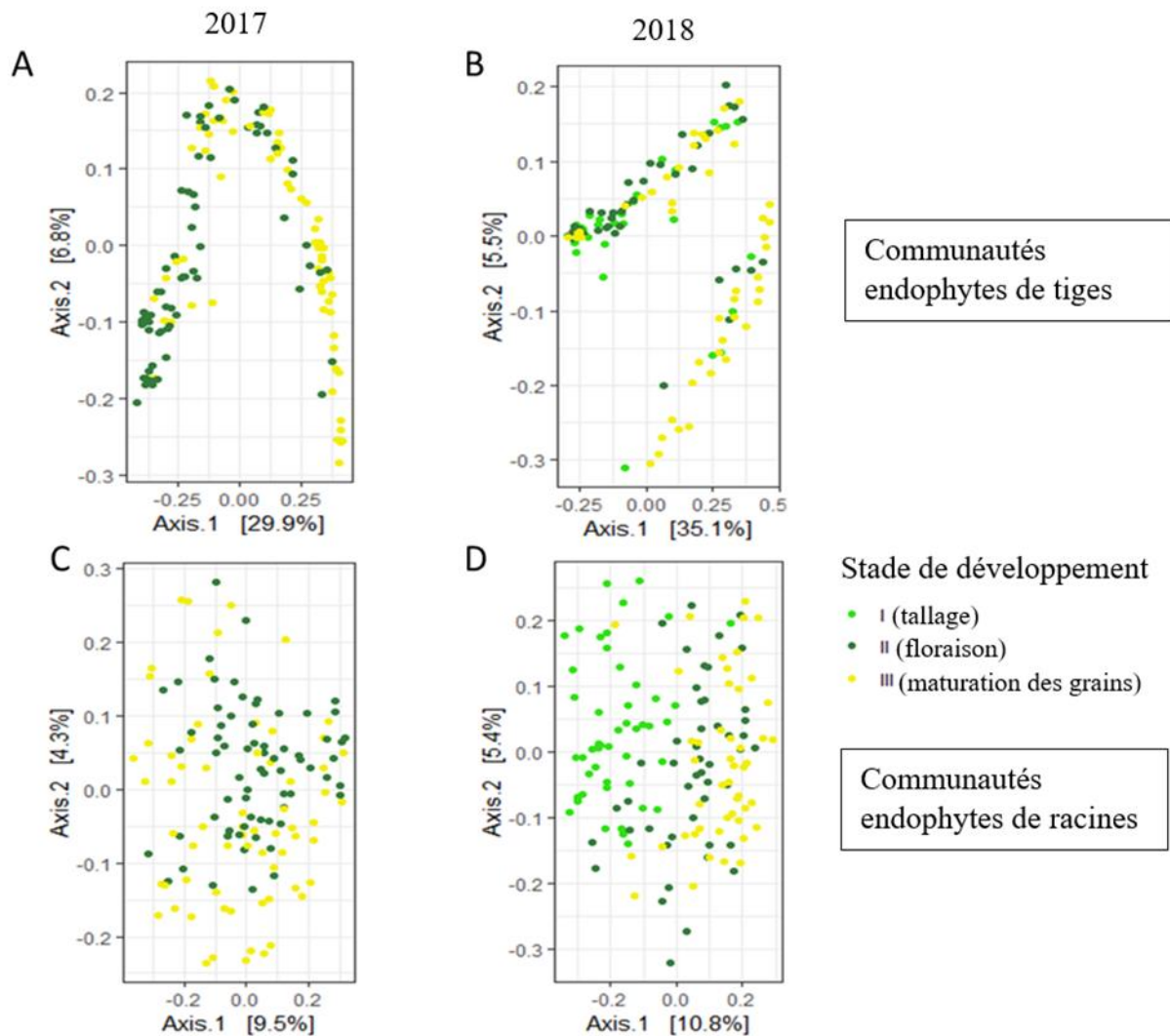


Figure 3.12 : Dynamique de structuration des communautés bactériennes endophytes des tiges et des racines de riz en fonction des stades de développement de la plante pour les années de prélèvement 2017 (A ; C) et 2018 (B ; D). Analyse en coordonnées principales (PCoA) sur l'indice UniFrac non pondéré. Les couleurs vert clair, vert foncé et jaune correspondent respectivement aux stades de développement du riz tallage, floraison et maturation des grains. (A ; B) analyse de la structure des communautés bactériennes endophytes des tiges et (C ; D) analyse de la structure des communautés bactériennes endophytes des racines de riz

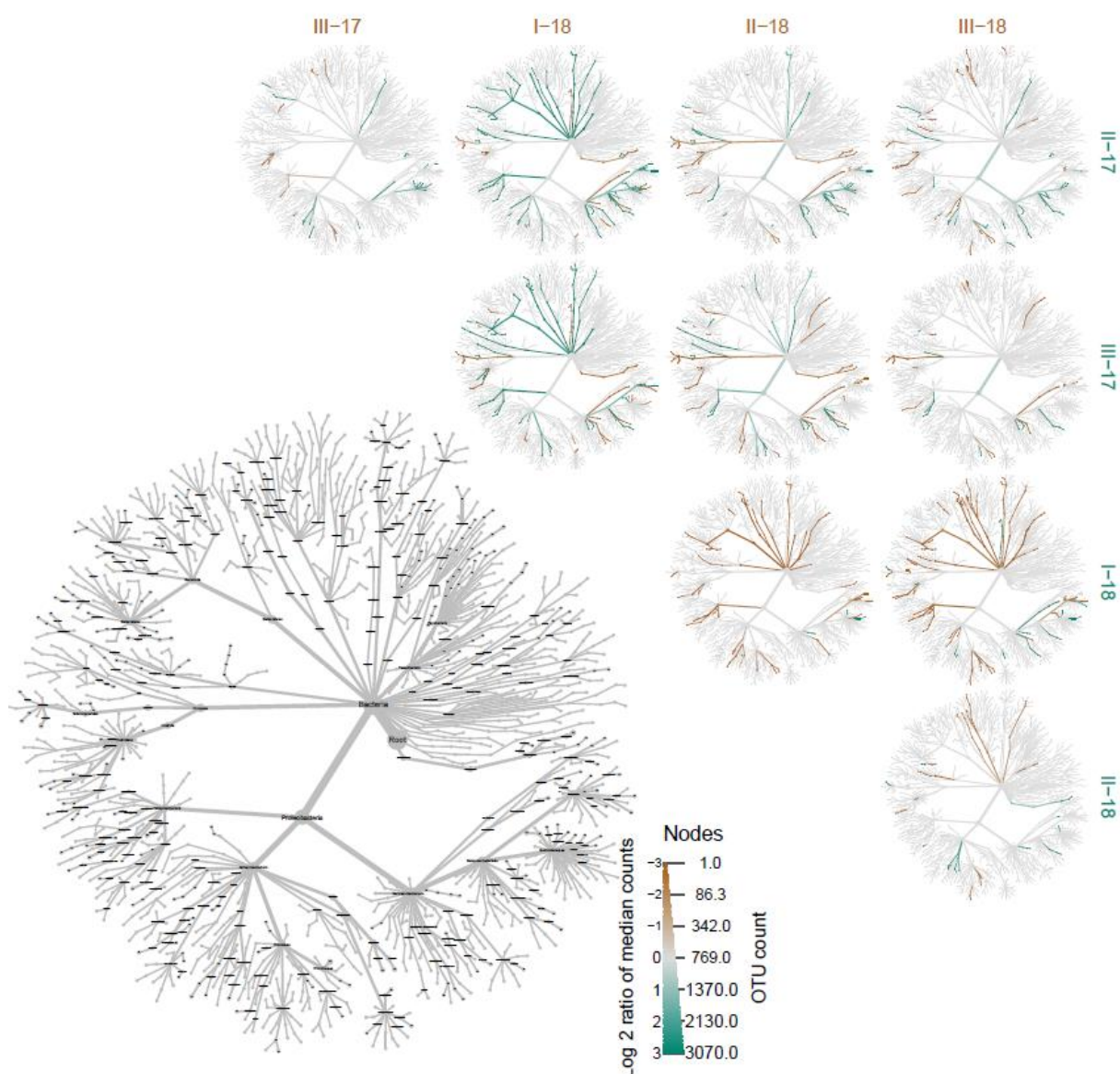


Figure 3.13 : Abondance différentielle des taxa endophytes des racines entre les stades de développement. L'arbre taxonomique gris représente la distribution taxonomique des bactéries endophytes du riz obtenue à partir de l'ensemble des échantillons analysés, la taille des nœuds correspond à l'abondance relative de chaque taxon. Cet arbre sert de clé pour identifier les taxa mis en évidence dans les « Heat tree différentiel » qui ne sont pas annotés. Les arbres taxonomiques sont appelés « Heat tree différentiel » ils représentent les résultats des comparaisons par paires de l'abondance des taxa bactériens entre les stades de développement du riz et entre les années de prélèvements. La couleur de chaque taxon représente le rapport log-2 de la médiane des abondances observées avec chaque groupe (racines, tiges et graines) et correspond aux différentes comparaisons par paire réalisée la couleur marron indique une abondance des taxa significativement (test de Wilcoxon, $p < 0,05$) plus élevée dans la variable positionnée en colonne lors de la comparaison par paire, la couleur verte indique une abondance des taxa significativement plus élevés dans la variable positionnée en ligne lors de la comparaison par paire. Tous les taxa de couleur grise ne présentent pas d'abondance relative significativement différente entre groupes comparés.

Les mêmes analyses de PERMANOVA réalisées sur les données de 2018, montrent un effet plus marqué du stade de développement (tallage, floraison et maturation des graines) sur la structure des communautés bactériennes endophytes des racines, avec 9% des variations qui sont expliquées par le stade de développement de la plante (Tableau 3-7). Concernant l'analyse de PCoA sur l'indice d'UniFrac non pondérée qui illustre les résultats de la PERMANOVA, nous pouvons voir une séparation plus nette sur le premier axe entre les communautés bactériennes endophytes des racines au stade tallage et celles aux stades floraison et maturation des graines (Figure 3-12 D). Les deux derniers stades de développement du riz (correspondant tous les deux au stade reproductif) sont moins distinctement séparés comme le montrent aussi les données de 2017 (Figure 3-12 C). Les résultats des PcoA suggèrent une évolution des communautés microbiennes endophytes des racines, plus marquée entre les stades végétatifs et reproductifs du cycle de développement du riz, qu'entre les stades floraison et maturation des grains du stade reproductif. La variation de structure des communautés microbiennes endophytes des racines, entre le stage végétatif et les stades reproductifs, n'a été estimée que sur une année de prélèvement. Les résultats supportent des travaux antérieurs réalisés sur le riz aux champs, qui ont montré une plus forte variation de la structure des communautés bactériennes associées aux racines de riz lors des stades végétatifs, puis une stabilisation de celles-ci à partir des stades de développement reproductifs (Edwards et al., 2018). D'autres travaux réalisés sur le riz en conditions contrôlées en serre, ont montré que l'assemblage du microbiote racinaire endophyte était initié dès le premier jour de contact avec les divers microbes du sol, et que cette acquisition était rapide avec une augmentation significative de la richesse bactérienne au cours des deux premières semaines (Edwards et al., 2015, Zhang et al., 2018). Ces trois études sur le riz qui ont mis en évidence des variations de structure des communautés bactériennes endophytes dans les racines au cours du temps, n'ont pas identifié la raison des différences de structures observées entre les stades de développement du riz. Des variations de composition des exsudats racinaires, au cours du cycle de développement du riz, pourraient expliquer des variations de structure des communautés microbiennes endophytes des racines : soit par des processus sélectifs de la plante soit par des processus de compétition où des microorganismes pourraient supplanter les existants (Aulakh et al., 2001). Des travaux sur *Arabis alpina* ont montré que le temps de résidence de la plante dans le sol et non le stade de développement, influençait les changements de structure du microbiote racinaire dans la structure des communautés bactériennes (Dombrowski et al., 2017). Nos résultats ne permettent pas de tester le temps de séjour dans le sol pour expliquer les variations de structures des communautés microbiennes associées au riz. Bien que cette variable soit probablement

confondue avec les cycles de développements du riz, nous ne pouvons seulement avancer dans nos travaux que le stade de développement du riz a vraisemblablement un rôle non négligeable dans les variations du microbiote observées.

A la suite de ces analyses des variations de structures des communautés bactériennes entre différents stades de développement du riz, nous avons comparé les variations d'abondances taxonomiques observées entre les stades de développements afin d'identifier des taxa bactériens dont l'abondance relative serait différenciellement et significativement associée à un ou plusieurs stades du cycle de croissance du riz (Figure 3-13). Ces analyses montrent de fortes variations d'abondances relatives, pour certains endophytes des racines, entre les stades tallage (I-18) et floraison (II-18) en 2018. Des variations d'abondances relatives significatives sont observées pour un nombre de taxa moins important, entre les stades floraison et maturation des grains (III-18) (Figure 3.13). Ces résultats suggèrent une phase de stabilisation des abondances relatives des taxa du microbiote endophyte des racines après la phase végétative. Ils indiquent à la fois une stabilisation qualitative (distance d'UniFrac non pondérée) et quantitative (abondances relatives) des communautés bactériennes endophytes des racines et confirment les résultats de précédents travaux sur la dynamique du microbiote racinaire du riz (Edwards et al., 2018, Zhang et al., 2018). Les variations qualitatives et quantitatives de la structure du microbiote racinaire, en fonction du stade de développement de riz, sont probablement associées aux variations de production des exsudats racinaires qui peuvent attirer différenciellement certains microbes (Bacilio-Jiménez et al., 2003, Berg & Smalla, 2009). La liste des taxa significativement et différenciellement abondants entre les stades de développement est présentée en Annexe 2.

- Dynamique de structuration des communautés bactériennes endophytes des tiges

Mes travaux de thèse apportent un regard complémentaire et nouveau aux travaux cités précédemment, qui ont tous focalisé sur le compartiment racinaire incluant la partie endophyte (Edwards et al., 2015, Edwards et al., 2018, Zhang et al., 2018) par l'étude de la dynamique de structuration des communautés bactériennes endophytes de la partie « aérienne de la plante » (base de la tige). Bien que les diagrammes en violon (Figure 3.11 A/B) semblent montrer à la fois en 2017 et en 2018, une augmentation de richesse bactérienne des communautés endophytes de la tige au stade de développement maturation des grains, aucune différence significative de richesse bactérienne n'a été observée entre les stades de développement du riz. Les analyses de PERMANOVA sur les distances de Jaccard et UniFrac non pondérées par

l'abondance des OTUs calculées entre les échantillons de tiges prélevés en 2017 (Tableau 3.7), indiquent un effet significatif des stades de développement sur la structure des communautés bactériennes avec environ 8% d'explication de la variance. L'analyse de PCoA sur la distance d'UniFrac non pondéré illustre bien les résultats de la PERMANOVA en montrant une séparation entre les communautés bactériennes endophytes des deux stades de développement échantillonnés (floraison à 70 jours après semis et maturation des grains, à 110 jours après semis) (Figure 3.12).

Les analyses de PERMANOVA réalisées entre les échantillons de tiges prélevées en 2018 à trois stades de développement, montrent que les structures des communautés bactériennes endophytes des tiges aux stades tallage, floraison et maturation des grains sont significativement différentes (Tableau 3.7). L'analyse PCoA suggère par ailleurs que les structures des communautés bactériennes endophytes des tiges aux stades tallage et floraison sont relativement semblables et sont toutes les deux différentes de la structure des communautés au stade maturation des grains (Figure 3.12).

En complément, une augmentation significative de l'abondance de plusieurs taxa est observée entre le stade floraison (II) et le stade maturation des grains (III) (Figure 3.14), mais aucune variation significative d'abondance d'OTUs n'est observée entre les stades tallage (I) et floraison (II). Ces résultats suggèrent à la fois une phase de stabilité qualitative (composition) et quantitative (abondances relatives) des communautés bactériennes endophyte des tiges entre le stade tallage (I) et le stade floraison (II) (Figure 3.12). Ces résultats indiquent par ailleurs une accumulation de certaines bactéries dans l'endosphère de la tige, entre le stade floraison (II) et le stade maturation des grains (III). La liste des taxa significativement et différenciellement abondants entre les stades de développement est présentée en Annexe 3. Aucun taxon bactérien n'est identifié comme significativement moins abondant au cours du temps.

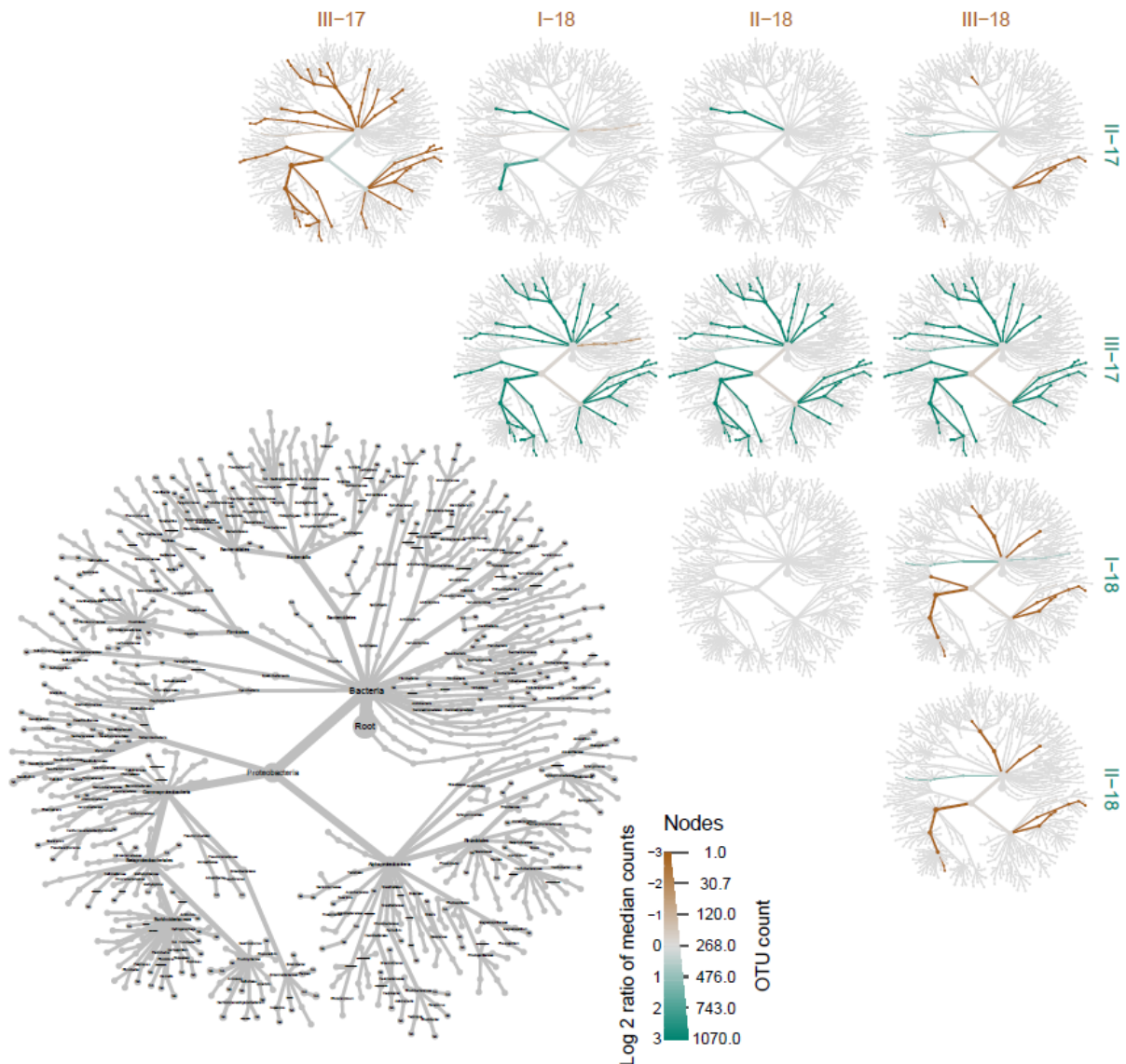


Figure 3.14 : Abondance différentielle des taxa endophytes des tiges entre les stades de développement. L'arbre taxonomique gris représente la distribution taxonomique des bactéries endophytes du riz obtenue à partir de l'ensemble des échantillons analysés, la taille des nœuds correspond à l'abondance relative de chaque taxon. Cet arbre sert de clé pour identifier les taxa mis en évidence dans les « Heat tree différentiel » qui ne sont pas annotés. **Les arbres taxonomiques sont appelés « Heat tree différentiel » ils représentent les résultats des comparaisons par paires de l'abondance des taxa bactériens entre les stades de développement du riz et entre les années de prélèvements.** La couleur de chaque taxon représente le rapport log-2 de la médiane des abondances observées avec chaque groupe (racines, tiges et graines) et correspond aux différentes comparaisons par paire réalisée la couleur marron indique une abondance des taxa significativement (test de Wilcoxon, $p < 0,05$) plus élevé pour la variable positionnée en colonne lors de la comparaison par paire, la couleur verte indique une abondance des taxa significativement plus élevés pour la variable positionnée en ligne lors de la comparaison par paire. Tous les taxa de couleur grise ne présentent pas d'abondance relative significativement différente entre groupes comparés.

- Conclusion sur la dynamique de structuration des communautés bactériennes endophytes du riz

En résumé, il a été montré dans des études précédentes que la diversité α des communautés microbiennes associées au riz impliquant à la fois richesse et abondance des taxa, changeait significativement dans les stades de développement précoces du riz, au cours des quinze premiers jours chez le microbiote endophyte des racines (Edwards et al., 2015). Nos travaux réalisés à 47, 70 et 110 jours après semis montrent que les variations de richesse bactérienne ne sont pas significativement différentes entre les stades de développement du riz à partir du stade tallage, et que les dynamiques de structuration des communautés bactériennes endophytes des racines et des tiges diffèrent suivant les stades de développement de la plante. En comparaison des communautés bactériennes des racines qui se stabilisent entre la floraison et la maturation des grains, les communautés bactériennes des tiges vont subir leurs principaux changements de composition à cette période du cycle de vie de la plante. Ces changements de structure de communautés bactériennes endophytes observés entre les stades de développements du riz, pourraient être le résultat, à richesse égale, d'équilibres relatifs entre phénomènes d'enrichissement des communautés par sélection active de la plante (gain de taxa phyllosphériques, colonisation par de nouveaux taxa racinaires, etc.) et de phénomènes de pertes (mécanismes de défense de la plante, compétition entre taxa, dérive génétique, etc.). Pour poursuivre ce raisonnement, nous avons analysé de façon croisée les tables d'OTUs des échantillons de racines et de tiges. Cette analyse montre que quelques OTUs sont exclusivement présentes dans la tige de riz et sont absentes des racines. Mais leurs prévalences dans les échantillons de tiges ne sont pas assez fortes pour observer des différences significatives d'abondance entre les tiges et les racines. La faible prévalence des OTUs retrouvées spécifiquement dans la tige souligne une certaine variabilité de structure des communautés bactériennes endophytes des parties aériennes et peut suggérer la colonisation par certains individus de bactéries à partir de la sphère épiphyte hétérogène des feuilles ou une colonisation hétérogène à partir des bactéries endophytes de racines (Redford et al., 2010, Bulgarelli et al., 2013). Par rapport aux deux autres stades de développements précédents, toutes les OTUs identifiées dans les échantillons de tiges comme relativement plus abondantes au stade maturation des grains, sont toujours présentes dans les communautés racinaires avec des abondances relatives toujours plus élevée dans les échantillons de racines que dans les échantillons de tiges, quel que soit le stade de développement du riz (Annexe 1). Cette observation suggère que les OTUs identifiées comme plus abondantes dans les tissus de la tige

de riz en fin de cycle de développement, sont des d'OTUs qui sont très probablement issues des racines de riz et qui ont diffusé dans les tiges en fin de cycle.

Nos résultats suggérèrent que les taxa initialement présents dans les racines se retrouvent à un stade plus tardif dans la tige avec un effet « retard » lié au temps de migration entre ces deux compartiments. Cette interprétation serait en accord avec des travaux précédents qui ont indiqué que le microbiote endophyte racinaire contribuerait à la colonisation de l'endosphère de la tige (Mano & Morisaki, 2008), dans une proportion qui reste à déterminer. Une autre hypothèse envisageable serait que le changement phénologique de la partie aérienne de la plante, avec son entrée en sénescence lors de la maturation des grains, entraîne des changements drastiques des disponibilités en nutriments pour les bactéries endophytes (Noodén et al., 1997, Zhang et al., 2010a), modifiant les pressions de sélection et entraînant le changement de composition des communautés bactériennes. On peut supposer que les deux hypothèses énoncées ci-dessus et probablement d'autres ne sont pas mutuellement exclusives et pourraient toutes contribuer en partie à expliquer les différences observées.

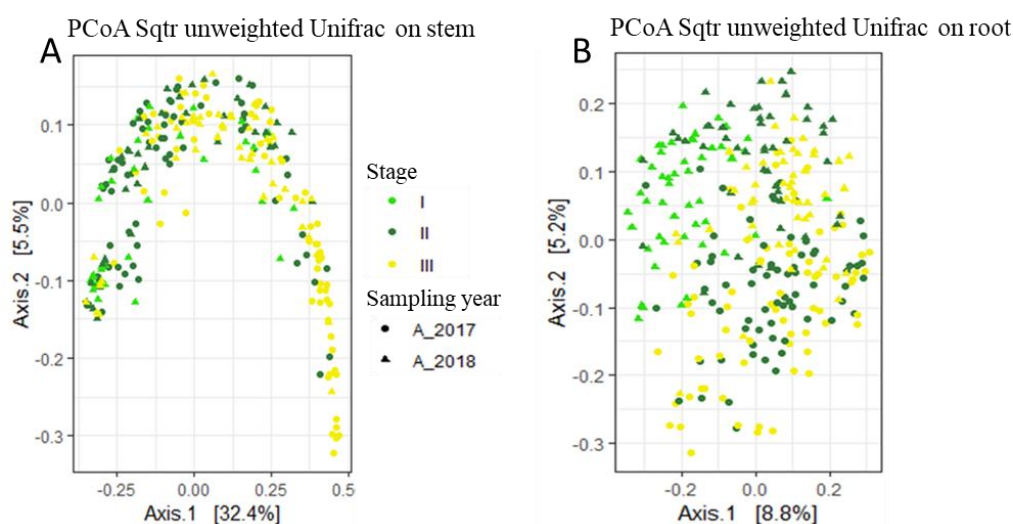


Figure 3.15 : Dynamique de structuration des communautés bactériennes endophytes des tiges et des racines de riz en fonction des stades de développements de la plante, Analyse en composantes principales (PCoA) sur l'indice UniFrac non pondéré. Les couleurs vert clair, vert foncé et jaune correspondent respectivement aux stades de développement du riz tallage, floraison et maturation des grains. Les formes des points du graphique (disque et triangle) correspondent aux années de prélèvements. (A) analyse de la structure des communautés bactériennes endophytes des tiges et (B) analyse de la structure des communautés bactériennes endophytes des racines de riz.

Cette dynamique de structuration des communautés bactériennes endophytes du riz s'est répétée en 2017 et 2018, ce qui suggère que le stade de développement de la plante est plus

structurant que les effets liés aux variations environnementales entre les deux années (Figure 3.15).

Nos résultats supportent l'hypothèse d'une colonisation des tissus racinaires par les microorganismes du sol, suivie d'une sélection de certains taxa par la plante qui s'établissent dans le compartiment racinaire endophyte et par la diffusion ou la migration de certains endophytes dans l'endosphère des tissus aériens (Copeland et al., 2015, Coleman-Derr et al., 2016, Bai et al., 2015, Compant et al., 2019, Compant et al., 2010). L'utilisation de plants axéniques (sans microorganisme associé) qui seraient transplantés dans un sol avec microbiote ou un substrat stérile à différents stades de développement, pourrait permettre d'apporter des éléments d'explications aux variations de structure des communautés bactériennes endophytes du riz dans le temps. Une analyse des communautés microbiennes du sol, aurait également permis d'appuyer l'hypothèse forte de la colonisation de la plante à partir de microorganismes du sol avec la mise en évidence d'un possible gradient. Cette analyse est en cours.

3.3.2.4. *Le génotype de la plante n'influence que modérément la structure des communautés bactériennes.*

Des effets significatifs du génotype de la plante sur la composition de son microbiote endophyte n'ont été observés que sur quelques comparaisons (floraison et maturation pour les racines en 2017 ; tallage et maturation pour les tiges en 2018) (Tableau 3.8). Aucun de ces effets n'est répétable entre les stades de développement et/ou entre les compartiments de la plante sur les deux années. Les analyses de PCoA sur distance d'UniFrac non pondéré confirment l'absence de structuration des communautés bactériennes endophytes par le génotype de la plante (annexe 4). Ces résultats sont en accord avec la littérature qui décrit des effets du génotype de la plante faibles à inexistantes sur la structuration des communautés microbiennes lorsque les comparaisons sont réalisées au sein de la même espèce de plante (Bulgarelli et al., 2012, Lundberg et al., 2012, Bouffaud et al., 2014, Peiffer et al., 2013, Bulgarelli et al., 2015, Schlaeppi et al., 2014). Les six variétés de riz analysées dans notre travail appartiennent toutes à la sous-espèce *Oryza sativa japonica* tempéré. La proximité génotypique de ces variétés peut expliquer l'absence d'effet marqué du génotype sur les communautés bactériennes endophytes. Nos résultats confirment les travaux d'Edwards et al (2018) sur la structure des communautés endophytes des racines chez trois variétés de riz appartenant toutes à la sous-espèce *Oryza sativa japonica* tempéré (Edwards et al., 2018). Dans ce travail, le génotype de la plante n'a expliqué qu'une mince partie des variations de structure du microbiote endophyte racinaire avec

moins de 1% d'explication de la variance. Dans une étude précédente analysant des variations du microbiote racinaire du riz chez six variétés appartenant à deux espèces différentes (*Oryza sativa japonica* et *Oryza glaberrima*), 1.5% de la variance était expliqué par le génotype de la plante (Edwards et al., 2015).

Tableau 3.8: Estimation de l'influence de la variété de riz sur la structure des communautés bactériennes endophytes du riz sur chaque prélèvement réalisé par des analyses de variance multivariée PERMANOVA basées sur l'indice de Jaccard et l'indice UniFrac non pondéré (10 000 permutations), sont représentés en gras les résultats significatifs avec une P-value inférieure à 0.05, la valeur R² désigne la proportion de la variance qui peut être expliquée par chacune des variables étudiées

			Jaccard		UniFrac non pondéré	
			R ²	Pr (< F)	R ²	Pr (< F)
Racine	2017	Floraison	0.092	3e-04	0.096	2e-04
		Maturation	0.087	0.021	0.082	0.142
	2018	Tallage	0.111	0.190	0.105	0.450
		Floraison	0.106	0.514	0.104	0.720
		Maturation	0.116	0.061	0.106	0.405
Tige	2017	Floraison	0.073	0.468	0.061	0.817
		Maturation	0.076	0.337	0.111	0.181
	2018	Tallage	0.147	0.014	0.156	0.029
		Floraison	0.117	0.205	0.120	0.200
		Maturation	0.140	0.004	0.154	0.003

3.3.2.5. Absence d'évidence d'un microbiote taxonomique cœur pour les communautés endophytes du riz

Un microbiote cœur du riz a été recherché en faisant l'hypothèse d'une sélection de taxa endophytes par la plante. Ce microbiote cœur prévalent dans la majorité des plantes, devrait contenir des taxa endophytes bénéfiques potentiellement essentiels au bon fonctionnement du riz. En considérant une prévalence supérieure à 80 % des taxa comme seuil d'appartenance au microbiote cœur, sans critère d'abondance, nous n'avons pas identifié de façon évidente un microbiote cœur pour les communautés bactériennes endophytes des tiges de riz. Aucune OTU n'a été retrouvée dans au moins 80% des échantillons, quelle que soit l'année de prélèvement. Il n'a pas non plus été observé de microbiote cœur pour un stade de développement spécifique. Pour les communautés bactériennes racinaires, seule une OTU endophyte attribuée au genre

Sulfurospirillum a été retrouvée associée à l'endosphère des racines de riz, avec une prévalence de plus de 80% pour toutes les dates d'échantillonnage (Tableau 2.9). Ce genre n'est pas connu pour avoir des effets bénéfiques pour la plante, mais a été décrit pour avoir un rôle dans la phytoremiadiation des sols dans les zones humides et dans les environnements en anaérobiose (Luijten et al., 2004, Li et al., 2010).

Nous avons remarqué que les racines chez les plantes au stade maturation des grains en 2017 ont une quasi-absence de microbiote cœur avec seulement 2 taxa avec plus de 80% de prévalence entre les échantillons (Tableau 2.9). Si nous ne prenons pas en compte ce point d'échantillonnage, le microbiote cœur des communautés bactériennes racinaires endophytes est augmenté de 3 OTUs présentes tout au long du cycle de croissance du riz (Tableau 2.9). Ces OTUs appartiennent aux genres *Azospirillum*, *Azospira*, et *Novosphingobium*, qui ont déjà été décrits pour voir des effets sur la croissance du riz (Bae et al., 2007, Drogue et al., 2014, Zhang et al., 2016). Par exemple *Azospirillum* est l'un des genres de rhizobactéries promotrices de la croissance des plantes les plus étudiés (Holguin et al., 1999, Steenhoudt & Vanderleyden, 2000, Bashan et al., 2004, Mano & Morisaki, 2008). Il est connu pour permettre la fixation l'azote atmosphérique (Bashan & de-Bashan, 2010), mais aussi pour sa capacité à stimuler la synthèse de phytohormones pouvant entre autres avoir un effet sur les structures racinaires du riz (Steenhoudt & Vanderleyden, 2000, Somers et al., 2004). Des bactéries du genre *Azospirillum*, ont déjà été retrouvées dans des associations sur riz en Camargue, avec l'isolement d'une souche d'*Azospirillum lipoferum* sur la variété de riz camarguaise Cigalon (Drogue et al., 2014).

Un microbiote endophyte cœur transitoire a été observé chez le riz contenant des OTUs bactériennes potentiellement fixatrices d'azote. Ainsi, trois OTUs sont retrouvées dans le microbiote cœur endophyte aux stades tallage et floraison et sont absentes au stade maturation des grains et cela de façon répétée sur les deux années de prélèvements (Tableau 2.9). Enfin plusieurs OTUs apparaissent dans le microbiote cœur endophyte des racines de façon non répétable entre les années.

Tableau 3.9: Liste des OTUs fortement conservées (prévalence supérieure à 80%) et considérées comme appartenant au microbiote cœur des racines de riz pour chaque stade de développement et années de prélèvements. Les croix indiquent pour chaque prélèvement la présence des OTUs avec une prévalence supérieure à 80%. Sont présentés seulement les taxa considérés comme appartenant au microbiote cœur présent dans au moins deux prélèvements. La liste complète des taxa des microbiotes cœur pour chaque prélèvement est présentée en annexe 5

Phylum	Class	Order	Family	Genus	Species	I 2018	II 2017	II 2018	III 2017	III I2018
Campilobacterota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	na	x	x	x	x	x
Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Azospirillum	na	x	x	x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Azospira	na	x	x	x		x
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	na	x	x	x		x
Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Niveispirillum	na	x	x	x		
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	Rhizobium_ipomoeae	x	x	x		
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Hydrogenophaga	na	x	x	x		
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax	na	x		x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Propionivibrio	na		x	x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	na	na		x			x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	na	na			x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax	na	x	x			
Desulfobacterota	Desulfovibrionia	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	na			x		x
Verrucomicrobiota	Verrucomicrobiae	Opitutales	Opitutaceae	Lacunisphaera	na		x			x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Candidatus_Accumulibacter	na			x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Propionivibrio	na	x				x
Spirochaetota	Spirochaetia	Spirochaetales	Spirochaetaceae	Spirochaeta_2	Spirochaeta_aurantia				x	x
Campilobacterota	Campylobacteria	Campylobacterales	Sulfurimonadaceae	Sulfuricurvum	na		x	x		

Le microbiote endophyte des grains de riz n'a présenté qu'une seule OTU partagée à plus de 80% dans les échantillons et appartenant au genre *Pantoea*. Ce genre, qui appartient à la famille des *Enterobacteriaceae*, est très largement décrit dans les études menées sur les semences de riz (Cottyn et al., 2009, Kaga et al., 2009, Zhang et al., 2019) car communément isolé (Mano & Morisaki, 2008) et présenté comme des taxa du microbiote cœur endophyte du riz le plus abondant (Cottyn et al., 2009). Un microbiote cœur des graines de riz exhibant 21 variants de séquences d'amplicons a été décrit, mais cela à partir de graines non désinfectées avec les enveloppes (Eyre et al., 2019). Plusieurs OTUs appartenant à des genres bactériens connus pour être associés plus ou moins intimement aux semences de riz (Cottyn et al., 2009, Zhang et al., 2019, Hardoim et al., 2012), ont été retrouvées dans nos données avec un seuil de prévalence inférieur à 80%. Il s'agit des genres *Pseudomonas* (68% de prévalence), *Xanthomonas* (32% de prévalence), *Microbacterium* (5,5% de prévalence), *Rhizobiaceae* (5,7% de prévalence) et *Sphingomonas* (4,2% de prévalence). Toutefois si nous ne portons pas attention à la prévalence, il est intéressant de noter que ces cinq genres bactériens sont connus pour faire partie du microbiote cœur des grains de riz ; ils représentent 98,5% des séquences assignées au genre dans nos données alors que le genre *Pantoea* seul ne représente que 64,8% des séquences.

Nos travaux qui suggèrent une très faible fraction d'OTUs appartenant au microbiote cœur du riz posant la question de la signification biologique du microbiote cœur taxonomique défini au niveau de l'OTU. L'analyse du microbiote cœur avec un clustering des séquences au niveau de l'OTU ne prend, en effet, pas en compte la redondance fonctionnelle et il est possible qu'un microbiote cœur fonctionnel du riz puisse être assuré par différentes communautés microbiennes (Tian et al., 2017). De plus, il est admis que la diversité du microbiote constitue un critère fort de résilience des écosystèmes face à des stress biotiques et abiotiques. Une diversité élevée est garante d'une redondance fonctionnelle qui permet à l'écosystème de maintenir en dépit de perturbations les fonctions importantes pour le complexe hôte/microbiote (Tian et al., 2017, Moya & Ferrer, 2016). Le cœur fonctionnel semble plus pertinent à estimer qu'un cœur taxonomique chez le microbiote des plantes (Lemanceau et al., 2017), mais les technologies de séquençage actuelles ne nous permettent pas d'estimer facilement la diversité fonctionnelle de l'holobionte.

3.3.3. Conclusions

Notre travail réalisé en conditions expérimentales sur des parcelles de riz en Camargue au cours de deux années visant à caractériser le microbiote endophyte chez *Oryza sativa Japonica* tempéré a révélé cinq conclusions principales :

- 1 L'organe de la plante structure majoritairement les communautés bactériennes endophytes.
- 2 Les communautés bactériennes endophytes du riz sont différentes en fonction de l'âge de la plante et sont probablement liées au stade phénologique de la plante.
- 3 Le microbiote taxonomique cœur endophytes du riz est très limité.
- 4 Le génotype des variétés de riz *japonica* tempéré ne structure que très peu les communautés bactériennes endophytes.
- 5 La présence d'un virus persistant du riz est structurée par le génotype du riz et n'a pas d'influence sur la structure des communautés bactériennes endophytes du riz et *vice versa*.

Discussion générale et perspectives

3.4. Rappel des objectifs de thèse

L'objectif principal de ce travail de thèse a été, dans le cadre général d'une meilleure compréhension de l'assemblage du microbiote des plantes, de déterminer si la structure des communautés microbiennes associées à une plante peut être influencée par le génotype de celle-ci. En parallèle de cet objectif principal, nous avons essayé de voir si la structure des communautés bactériennes et fongiques était significativement modifiée par la présence/absence de phytovirus au sein de plants de riz collectés en Chine et en Camargue.

Nous avons dans ce cadre général, cherché à caractériser trois composantes majeures du microbiote du riz, à savoir : les virus, les bactéries et les champignons. A notre connaissance, ce travail de recherche est le seul à avoir pris en compte à la fois le compartiment endophyte aérien et souterrain du riz et à avoir inclus, dans les travaux de caractérisation du microbiote, la diversité virale en plus de la diversité bactérienne et fongique.

3.5. Caractérisation des microbiotes du riz

3.5.1. Une dynamique dans l'espace

Dans les deux chapitres de ce travail de thèse, nous avons identifié le compartiment de la plante comme étant le facteur majeur de structuration des communautés microbiennes du riz. Ce résultat est en accord avec les conclusions formulées dans plusieurs autres études qui ont aussi montré un effet majeur du compartiment végétal (souterrain vs. aérien ou épiphytes vs. endophytes) dans la structuration des communautés microbiennes des plantes (Wallace et al., 2018, Knief et al., 2012, Bertani et al., 2016, Wang et al., 2016, Hamonts et al., 2018). Les variations environnementales ou de niches sont largement reconnues comme des facteurs impactant les processus d'assemblage des communautés microbiennes (Cregger et al., 2018, Singer et al., 2019). Les conditions physicochimiques proposées dans chaque niche écologique des différents compartiments de la plante génèrent, en effet, une pression de sélection propre opérant sur les communautés microbiennes occupant ces compartiments. A titre d'exemple, les conditions environnementales diffèrent fortement entre les compartiments racinaires incluant sol, rhizosphère, rhizoplan et endosphère racinaire (Reinhold-Hurek et al., 2015), mais aussi entre les compartiments extérieurs de la plante, aériens au niveau de la phyllosphère ou souterrains au niveau de la rhizosphère et du rhizoplan (Vorholt, 2012, Müller et al., 2016). Les fortes variations d'environnement entre les compartiments végétaux rencontrées chez le riz en

conditions inondées, entraînent une sélection des microorganismes capables de s'adapter aux différentes conditions proposées. Les sources de nutriments, la présence ou l'absence d'oxygène, les conditions d'humidité, les radiations UV et de nombreux autres facteurs influencent les différentes communautés associées au riz. Nous confirmons l'existence de microbiotes distincts dans les compartiments externes aérien et souterrain chez le riz dans les terrasses du Yuanyang (Edwards et al., 2015, Edwards et al., 2018, Imchen et al., 2019, Knief et al., 2012). Nous montrons aussi que les deux compartiments endophytes au niveau des racines et de la tige ont des structures de communautés microbiennes différentes avec une diminution de richesse dans la tige. Si ces deux compartiments sont connectés par les tissus du parenchyme et les vaisseaux conducteurs, des pressions de sélection différente opérant sur ces compartiments, auxquelles s'ajoutent possiblement des événements stochastiques associés aux événements de migration/dispersion contribuant au microbiote endophyte aérien, peuvent expliquer les différences entre ces microbiotes.

3.5.2. Une dynamique dans le temps

L'étude de la dynamique du microbiote chez plusieurs génotypes de riz nous a indiqué que la richesse (nombre de bactéries différentes) des communautés endophytes était relativement stable du stade tallage au stade maturation des grains, comme chez *Arabidopsis thaliana* où la richesse du microbiote de la rhizosphère ne varie pas entre les stades plantule et floraison (Chaparro et al., 2014). En revanche, nous montrons que la composition des communautés bactériennes change en fonction des stades de développement du riz confirmant des travaux précédents effectués sur cette plante (Edwards et al., 2015, Edwards et al., 2018, Imchen et al., 2019, Zhang et al., 2018) et sur d'autres plantes (Chaparro et al., 2014). Plusieurs hypothèses peuvent expliquer ces variations temporelles. La première hypothèse s'appuie sur les changements d'état physiologique de la plante au cours de son développement (Chaparro et al., 2014, Bais et al., 2006). Comme pour les compartiments de la plante, des niches physicochimiques différentes peuvent être proposées par la plante au cours de son développement. Par conséquent, des changements rapides des conditions physicochimiques liés au développement de la plante peuvent engendrer des variations de la composition microbienne dans chaque compartiment. Dans ce cas, c'est la plante qui sélectionne le microbiote. La seconde hypothèse repose sur le phénomène de succession des communautés microbiennes c'est-à-dire sur les changements de communautés qui apparaissent suivant la colonisation d'un nouvel environnement stérile par les microorganismes (Green et al., 2006, Fierer et al., 2010).

Des espèces de bactéries pionnières vont coloniser en premier, les tissus de la plante. Puis les changements physicochimiques générés par ces microorganismes pionniers pourront impacter de façon positive ou négative, les autres espèces arrivant plus tard : ce sont les effets de priorité (Zhou & Ning, 2017). Au cours de ce mécanisme de succession contribuant à la structure des communautés, des processus stochastiques et déterministes façonnent cette succession avec ici, une pression de sélection exercée par les microorganismes. Les deux hypothèses énoncées ci-dessus ne sont probablement pas mutuellement exclusives et s'additionnent tout aussi probablement à d'autres processus stochastiques et déterministes impliqués dans l'assemblage et la persistance des communautés microbiennes tout au long du cycle de vie de l'hôte végétal.

Nos travaux apportent une nette avancée dans les connaissances sur la dynamique de structuration des communautés microbiennes du riz. Jusqu'à ce jour, les travaux sur ce thème et sur l'ensemble du cycle de développement du riz avaient porté essentiellement sur les compartiments racinaires (Edwards et al., 2015, Edwards et al., 2018, Imchen et al., 2019, Zhang et al., 2018). Par nos travaux en Camargue, nous avons analysé, pour la première fois sur la même plante, la dynamique de structuration des communautés microbiennes endophytes du riz, sur à la fois les compartiments végétaux aérien et souterrain. La dynamique de structuration des communautés microbiennes endophytes racinaires confirme la littérature (Edwards et al., 2015, Edwards et al., 2018, Imchen et al., 2019, Zhang et al., 2018). Nous montrons aussi un changement de composition des communautés bactériennes endophytes racinaires entre les stades de développement végétatifs et reproductifs. La composition des communautés bactériennes endophytes racinaires reste relativement stable tout au long du stade reproductif (entre la floraison et la maturation des grains). Edwards et coll. (2015) a suggéré que les microbiotes des trois compartiments racinaires, rhizosphère, rhizoplan et endosphère, étaient connectés. La composition de ces trois microbiotes distincts se chevaucherait en se déplaçant des sections les plus externes vers l'endosphère racinaire, chacun de ces compartiments partagerait un sous-ensemble des communautés. Une partie des microorganismes du sol coloniserait donc progressivement les différents compartiments de la rhizosphère à l'endosphère. Nous montrons par nos travaux, qu'une partie importante des bactéries présentes dans l'endosphère de la tige est également présente dans l'endosphère des racines, mais avec des abondances plus faibles. Dans ce modèle à plusieurs étapes, nos résultats suggèrent que le compartiment endophyte de la partie aérienne est distinct de celui des racines, bien que probablement issu de ce dernier. Certains microorganismes endophytes des racines seraient donc sélectionnés par la plante et continueraient de progresser dans celle-ci. Nous

montrons également que la dynamique de structuration des communautés bactériennes endophytes des tiges est différente de celle des racines. Alors que les communautés bactériennes des racines changent de structure essentiellement entre les stades végétatifs et reproductifs, la structure des communautés bactériennes endophytes des tiges n'évolue que très peu jusqu'à la floraison du riz. Un changement plus important est alors observé lorsque la plante, au stade maturation des graines commence à entrer en sénescence. Ces profils différents, suggèrent un « effet retard » dans la dynamique de l'assemblage entre les compartiments endophytes, aériens et souterrains.

Par ce travail, nous démontrons que le microbiote du riz est dynamique dans le temps et dans l'espace. La plante a la capacité de sélectionner, de filtrer et de favoriser les microorganismes associés spécifiquement à chacune des parties de la plante, au cours de son développement. Il est important dans les études de caractérisation du microbiote du riz de prendre en compte ces changements et de bien les comprendre pour éventuellement pouvoir les prédire. Dans nos travaux sur l'agrosystème rizicole en terrasses du Yuanyang, il serait intéressant de faire un suivi de la dynamique des communautés microbiennes, afin de caractériser la succession de ces communautés microbiennes au cours du cycle du riz. Ces données, associées à une connaissance de l'historique de la parcelle et à une caractérisation du microbiote du sol, pourraient aider à mieux appréhender la distribution spatio-temporelle de la diversité du microbiote du riz dans les terrasses, et mieux estimer son rôle potentiel dans la durabilité du système agricole.

3.5.3. Influence du génotype de la plante hôte sur la structure des communautés microbiennes associées au riz

Nous avons confirmé que le compartiment de la plante et son stade phénologique impactent la structure du microbiote du riz, dans les conditions de culture suivies dans la province du Yunnan et en Camargue. D'autres facteurs biotiques associés à l'hôte ou à son microbiote ou des facteurs abiotiques, peuvent impacter le microbiote des plantes (Compant et al., 2019). Par exemple, nous avons discuté le rôle des exsudats racinaires dans l'assemblage du microbiote racinaire et son évolution dans le temps. L'exsudation racinaire est génétiquement programmée et donc dépendante de la variabilité génétique de son hôte (Rengel, 2002). Nous avons testé l'influence du génotype de l'hôte sur la structure des communautés microbiennes associées au riz.

Identifier si le génotype de l'hôte contrôle une partie de la composition de ses communautés microbiennes, est un sujet de recherche actuellement très développé (Imchen et al., 2019, Andreo-Jimenez et al., 2019, Shenton et al., 2016, Shi et al., 2019, Walitang et al., 2018). Une finalité appliquée à ce type de recherche, serait de partiellement manipuler le microbiote à partir du choix variétal, pour répondre à des stress biotiques et abiotiques (Borbet & Blaser, 2019, Liu et al., 2018, Kreznar et al., 2017, Walitang et al., 2018).

Dans nos travaux, nous avons comparé les structures des communautés microbiennes associées à des génotypes de riz *indica* en Chine, en comparant deux groupes de variétés traditionnelles et modernes génétiquement éloignés dans la phylogénie mondiale du riz, et des génotypes de riz *japonica* tempérés en Camargue montrant une variabilité génétique plus étroite. Nous avons révélé dans les travaux des chapitres 1 et 2 que le génotype de la plante a un effet significatif, mais modéré à faible (1,5-6%, p-value < 0,05) sur la structure des communautés microbiennes retrouvées sur et dans les racines et les tiges de riz. Plus précisément, nous estimons l'effet du génotype entre 3 et 6% des variations des communautés épiphytes du riz, et 1,5% des variations des communautés bactériennes endophytes. Ces niveaux relativement faibles des effets du génotype de l'hôte sur les communautés microbiennes du riz sont en accord avec les résultats de plusieurs travaux de recherche sur le microbiote des plantes (Bulgarelli et al., 2013, Bulgarelli et al., 2015) dont certains ont été réalisés sur le riz (Andreo-Jimenez et al., 2019, Shenton et al., 2016, Shi et al., 2019, Edwards et al., 2015, Edwards et al., 2018). Ces études récentes, ont en effet révélé que le génotype du riz n'expliquait que faiblement (1-2,5%) la variance de la structure des communautés microbiennes racinaires chez le riz (Edwards et al., 2015, Edwards et al., 2018). Il est à noter cependant, que de faibles variations observées entre les génotypes de riz peuvent participer à des phénomènes de régulation, d'adaptation et de résistance de la plante, à des stress biotiques et abiotiques (Andreo-Jimenez et al., 2019). Par exemple, le génotype de l'hôte pourrait agir sur quelques taxa qui sont des « hubs » microbiens ou des taxa clés de voute interconnectés dans le réseau microbien ou importants pour sa structure (Agler et al., 2016).

De fait, alors que nos travaux et ceux d'Edwards et coll. (2015, 2018) ont été réalisés en l'absence de stress abiotique connu, des travaux complémentaires récents ont été menés sur le riz en condition de sécheresse et ont montré que dans ces conditions de stress le génotype de l'hôte impactait plus fortement la structure des communautés fongiques (Andreo-Jimenez et al., 2019). Toutefois, le pourcentage de variance expliquée dépendait de la prise en compte de l'abondance des OTUs dans le modèle de cette étude, qui testait l'effet du génotype de l'hôte

et du traitement « sécheresse » sur la structure des communautés fongiques associées aux racines de riz (Andreo-Jimenez et al., 2019). Ainsi, le génotype de l'hôte expliquait significativement 1% de la variance observée en ne tenant pas compte de l'abondance des OTUs, alors que cet effet passait à 13% en tenant compte de l'abondance des OTUs. Il est intéressant de noter que nos travaux ainsi que ceux d'Edwards et Coll. ne relèvent pas d'augmentation de l'effet du génotype du riz sur la structure des communautés microbiennes (bactéries et champignons) en considérant l'abondance des OTUs dans les calculs de diversités. L'ensemble de ces résultats suggère, qu'en condition de stress, les différents taxa microbiens sont très probablement différenciellement affectés. Seuls quelques microbes peuvent probablement supporter les conditions extrêmes du stress (ici la sécheresse) et se multiplier. Les microbes non adaptés à ces conditions peuvent rester en dormance en attendant de retrouver des conditions favorables à leurs croissances ou voir leurs populations diminuer. Dans cette logique, les microbes susceptibles de supporter les nouvelles conditions physicochimiques générées par le stress sur la plante ou sur le microbiote pourront avoir une abondance plus élevée. Par ailleurs, il est possible que la plante essaye de recruter de nouvelles fonctions pour répondre à ce stress et maximise ses relations avec les microorganismes bénéfiques, capables de l'aider à supporter le stress. Il est aussi possible que la plante et son environnement proche servent de refuge à certains microorganismes qui profitent d'une atténuation du phénomène de stress dans la sphère racinaire, où ils peuvent probablement mieux se développer. D'autres études que celle d'Andreo-Jimenez et coll. ont confirmé l'accentuation des différences de structures des communautés microbiennes chez le riz, sous des conditions de stress hydrique ou salin par rapport à des conditions de cultures sans stress connu (Andreo-Jimenez et al., 2019, Walitang et al., 2018, Santos-Medellín et al., 2017, Edwards et al., 2018).

Ces travaux effectués en conditions de stress sont probablement une piste intéressante à suivre pour évaluer pleinement les effets du génotype de l'hôte sur la structure de leur microbiote et anticiper des épisodes de stress abiotique en connaissant les capacités d'adaptation de l'holobionte riz. Par exemple, il serait intéressant d'analyser l'impact du génotype du riz à des stades clés de la colonisation microbienne, lors de la germination ou en situation de stress biotiques et abiotiques (par exemple des différences de salinité en Camargue) pour mieux évaluer la capacité de chaque génotype à mobiliser spécifiquement son microbiote. Ce type de travaux pourrait révéler comment différents génotypes de l'hôte peuvent exprimer leurs capacités à s'adapter en recrutant certains microorganismes spécifiques. L'étude du microbiote des plantes en conditions de stress a, en effet, dans plusieurs cas permis d'identifier des taxa

associés aux plantes impliquées dans des résistances à des stress biotiques (Berendsen et al., 2012, Berg, 2009, Mendes et al., 2013, Pieterse et al., 2014). Par exemple, l'utilisation de bactéries endophytes indigènes sélectionnées associées au jonc, augmente l'élimination de polluants organiques et métalliques (Syranidou et al., 2016). De même, des souches bactériennes isolées et caractérisées comme résistantes à des métaux lourds à partir de *Prosopis juliflora* poussant sur des effluents contaminés de tanneries ont été inoculés à du ray-grass et ont augmenté sa croissance et sa capacité à éliminer des métaux lourds sur ces mêmes effluents (Khan et al., 2015). Dans un autre exemple ciblant le biocontrôle, la colonisation d'*Arabidopsis thaliana* par certaines rhizobactéries déclenche la résistance systémique acquise vis à vis d'un panel d'agents pathogènes qui est médiée par le facteur de transcription MYB72, activé dans les racines (Van der Ent et al., 2008). Sous l'effet de MYB72, une coumarine ayant un effet antimicrobien, la scopoletine, est excrétée dans la rhizosphère et inhibe *in vitro* des champignons pathogènes, mais pas les rhizobactéries bénéfiques induisant la résistance systémique acquise (Stringlis et al., 2018). L'utilisation de telles souches permettrait d'améliorer la croissance et la protection de la plante.

3.5.4. Caractérisation du virome du riz et étude de son influence sur la composition des communautés microbiennes du riz

Le travail de métagénomique virale que nous avons réalisé, a apporté une première vision du virome du riz dans deux agrosystèmes. Comme pour d'autres cultures agricoles, l'approche de métagénomique virale sans à priori VANA s'est révélée être un bon outil de diagnostic pour détecter des virus à partir de plantes visuellement asymptomatiques (Palanga et al., 2016). L'approche de métagénomique virale VANA, nous a permis d'identifier de façon très inattendue deux virus du riz, avec dans les deux cas, une prévalence élevée. La découverte d'une situation épidémique d'un virus émergent dans l'agrosystème chinois centenaire et durable du Yuanyang est importante à considérer pour l'avenir de cet agrosystème. Le deuxième virus est *oryza sativa* alphaendornavirus, qui est connu pour être transmis verticalement, et qui n'avait jamais encore été décrit en Europe (Horiuchi et al., 2003). Sa présence au sein de variétés européennes suppose l'introduction et l'utilisation en Europe de semences contaminées par le biais de programmes de création variétale.

Les étapes techniques de la métagénomique virale VANA (François et al., 2018) ainsi que les étapes de bioanalyse comme l'interrogation des bases internationales dans l'assignation taxonomique des séquences virales, restent cependant perfectibles. Côté biologie moléculaire,

l'approche VANA, basée sur une semi-purification des particules virales, ne permet probablement qu'imparfaitement de décrire le virome du riz. A titre d'exemple, les bactériophages potentiellement associés aux bactéries du riz et régulateurs de leurs populations sont probablement mal caractérisés car ils ont, pour la majorité d'entre eux, une taille trop importante et doivent être systématiquement écartés au niveau de l'étape de filtration. Côté bioanalyse, la faible présence de séquences de phytovirus et de mycovirus dans les banques de données internationales rend difficile l'assignation taxonomique des séquences virales générées par les approches de métagénomique. Nous avons par exemple, mis en évidence la présence de plusieurs mycovirus dans les échantillons de riz, avec un possible rôle de biocontrôle des champignons pathogènes du riz (de la Paz Giménez-Pecchi et al., 2002, Yadav et al., 2015, Moriyama et al., 2018). Cependant, les identités nucléotidiques des séquences produites dans nos travaux avec les séquences des mycovirus référencées dans la base de données internationale GenBank étaient très faibles, suggérant que les mycovirus des rizières du Yunnan et de Camargue n'ont encore jamais été décrits. Cela réduit la portée de nos analyses en étant incapable de relier clairement ces virus à un hôte fongique identifié. Pour les mêmes raisons, il est fort probable que plusieurs séquences virales n'aient pas été identifiées par les approches d'interrogation des bases de données internationales et qu'une partie des séquences générées soit restée « muette », limitant *de facto* la profondeur de notre inventaire du virome du riz. Ces données de séquençage non assignées à des séquences connues est appelé par les anglo-saxons la « dark matter » (Roossinck et al., 2015). Une analyse précise de la diversité virale des virus de champignons et des virus de bactéries pourrait permettre de mieux comprendre les dynamiques de microorganismes bactériens et fongiques associées au riz. La compréhension des interactions microorganismes/microorganismes permettrait d'envisager des solutions de lutte biologique pour contrôler les microorganismes pathogènes du riz.

Suite à la découverte de la forte prévalence de deux phytovirus en Chine et en Camargue, nous avons voulu étudier l'influence de ces deux virus sur la structure des communautés microbiennes associées au riz. Ces deux virus ont des relations de symbioses différentes. Le premier est non persistant et pathogène pour la plante (« acute virus » pour les anglo-saxons) : il s'agit du Southern rice black streaked dwarf virus (SRBSDV) isolé dans les terrasses du Yuanyang qui est transmis de façon horizontale par une cicadelle (Alonso et al., 2019). Le second virus, oryza sativa alphaendornavirus, que nous avons retrouvé en Camargue est persistant et ne provoque pas de symptômes apparents sur la plante et est transmis de façon verticale par les semences de riz. De façon surprenante, nous n'avons pas détecté d'effet

significatif de la présence de chacun des deux virus sur la structure des communautés microbiennes associées au riz.

A notre connaissance, aucune étude sur plantes n'a à ce jour étudié l'effet d'une maladie virale sur le microbiote bactérien et fongique associé aux plantes. L'absence d'effet sur le microbiote du riz d'une maladie virale est un résultat très surprenant. Néanmoins, il est possible dans le cas de l'infection par le SRBSDV, que la détection très précoce du virus avant l'apparition de symptômes explique l'absence d'une modification de la structure du microbiote du riz. Il n'est pas exclu que des variations de structure des compositions microbiennes soient visibles à un stade plus avancé de la maladie. En effet, les changements du statut immunitaire de la plante, suite à la réponse contre l'agression virale ainsi que les lésions causées par le virus, pourraient entraîner des perturbations du microbiote. Il est par exemple connu chez l'Homme que les virus puissent avoir un effet bénéfique ou délétère sur les autres membres du microbiote, comme par exemple les bactéries (Rowan-Nash et al., 2019). Il a ainsi été montré que *Pseudomonas aeruginosa* peut utiliser les bactériophages filamenteux pendant la durée de vie des biofilms bactériens, ce qui améliore à la fois l'adhésivité des biofilms et leur protection contre la dessiccation et les antibiotiques (Secor et al., 2015).

Le cas de l'*Oryza sativa* alphaendornavirus (OsEV) est différent. Le riz semble le tolérer et le maintenir à une faible concentration dans la plante grâce à l'activation de la machinerie cellulaire du RNA silencing (Fukuhara, 2019). Ce virus n'a pas d'effet visible sur la fitness de la plante (Fukuhara, 2019) et ne semble pas perturber son fonctionnement et les échanges avec les microorganismes associés au riz, tout au long du cycle de vie de la plante. Dans la mesure de nos connaissances actuelles, ce virus semble avoir une relation de commensalisme avec le riz, il profite de la plante sans lui nuire. Dans de futurs travaux, il serait intéressant d'estimer les coûts et les bénéfices que pourrait représenter le maintien de ce virus à faible concentration pour la plante (Honjo et al., 2020, Roossinck, 2011). Des travaux menés par des équipes japonaises ont commencé de répondre à cette question, en ne montrant aucune différence de fitness de la plante à partir de lignées de riz infectées et non infectées par OsEV (Moriyama et al., 1996). Cependant d'autres travaux menés sur un endornavirus du haricot et prenant en compte plusieurs autres paramètres de la fitness de la plante, ont montré un effet positif du virus sur la fitness de la plante, comme par exemple une germination plus rapide des graines, des racines plus longues, des gousses plus longues, etc (Khankhum & Valverde, 2018). Il serait aussi intéressant de tester l'hypothèse d'un effet d'OsEV sur l'infection du riz par d'autres agents pathogènes ou sur les capacités de résistance de la plante à d'autres stress biotiques et

abiotiques. Il a par exemple été montré, que dans certaines conditions de stress abiotiques, l'infection de la plante par un virus augmente la survie de la plante (Gorovits et al., 2019).

Dans ces travaux, nous n'avons testé l'effet potentiel de deux phytovirus que sur les compositions microbiennes associées au riz. Il serait intéressant d'étendre ce travail en testant l'effet des mycovirus ou des bactériophages sur la composition du microbiote et le contrôle de la présence de microorganismes pathogènes, comme déjà rapporté dans d'autres études (Moriyama et al., 2018, Urayama et al., 2010, Zheng et al., 2019). Ces deux groupes de virus ont été mal estimés dans nos travaux, car mal purifiés et mal connus des bases de données. Il y a donc là l'opportunité d'affiner notre compréhension des interactions microbe-microbe bénéfiques à la santé de la plante hôte. D'autres études d'interaction microbe/microbe agissant sur le contrôle des microorganismes pathogènes sont par ailleurs possibles. L'occupation de la niche écologique par certains microorganismes peut éviter l'installation de communautés pathogènes (Pandin et al., 2017). Certains microorganismes peuvent aussi avoir un effet antagoniste sur la présence d'agents pathogènes (Xu et al., 2017, Neha et al., 2016). L'analyse de réseaux d'interaction entre microorganismes à la suite de nos travaux permettrait peut-être d'évaluer plus en profondeur l'importance des interactions microbes/microbes et d'explorer les capacités de complexes microbiens à affecter le statut physiologique de la plante (Layeghifard et al., 2017, Vacher et al., 2016b).

3.5.5. Processus déterministes et stochastiques façonnant les communautés microbiennes du riz

L'assemblage des communautés, incluant les microbes, est gouverné par quatre processus éco-évolutifs fondamentaux dont deux sont responsables de l'apport de nouveaux organismes dans la communauté, i.e. la dispersion et la diversification, et deux autres régulent l'abondance relative des espèces, i.e. la sélection et la dérive écologique (Zhou & Ning, 2017, Cordovez et al., 2019, Vellend, 2010, Nemergut et al., 2013). Deux de ces processus ont une composante uniquement déterministe (sélection) ou stochastique (dérive écologique). La dispersion et la diversification sont surtout des processus stochastiques, même si une composante déterministe pourrait aussi contribuer à ces deux processus (Zhou & Ning, 2017). Cependant la contribution relative de ces différents processus de nature déterministe et/ou stochastique reste difficile à estimer.

Chez le microbiote du riz, divers facteurs déterministes contribuent à son assemblage en exerçant différentes pressions de sélection, telles que des variations de conditions d'habitat proposées par les différents compartiments de la plante hôte (Edwards et al., 2015, Wang et al., 2016), de conditions environnementales (Edwards et al., 2015, Walitang et al., 2018, Santos-Medellín et al., 2017, Andreo-Jimenez et al., 2019) et de diversité génétique des plantes (Edwards et al., 2015, Andreo-Jimenez et al., 2019). Cependant, des mesures approfondies de nombreuses variables environnementales dans différentes études sur les écosystèmes associés aux plantes (Graham et al., 2016, Stegen et al., 2012, Ramette & Tiedje, 2007, Zhou et al., 2008, Hamonts et al., 2018) ont révélé qu'une part importante des variations dans la communauté microbienne ne pouvait pas être expliquée seulement par des facteurs déterministes.

Dans nos travaux, nous avons mis en place des designs expérimentaux à des échelles spatiales très resserrées en faisant l'hypothèse d'une homogénéité des facteurs biotiques et abiotiques. Ce postulat de base nous a orienté vers l'estimation de l'effet de facteurs associés à l'hôte comme ses compartiments, ses stades phénologiques ou son génotype. Nous avons montré que le compartiment et le stade phénologique associé à la physiologie de la plante expliquent en partie les variations spatio-temporelles observées chez le microbiote du riz alors que le génotype de l'hôte contribue quant à lui très peu à la structuration du microbiote. Ces facteurs font intervenir des mécanismes de sélection et sont donc essentiellement de nature déterministe même si des facteurs stochastiques peuvent intervenir lors du processus de succession contribuant aussi à la structuration du microbiote aux différents stades phénologiques étudiés (Cordovez et al., 2019, Dini-Andreote et al., 2015). D'autres facteurs déterministes non liés à l'hôte peuvent aussi contribuer à la structuration du microbiote du riz comme les facteurs environnementaux (climatiques, stress abiotiques, ...) ou liés aux interactions microbiennes (compétition, mutualisme ...) (Niu et al., 2017b). Dans nos travaux, la présence de deux virus de plantes, l'un pathogène et l'autre commensal, n'a cependant pas altéré le microbiote du riz. Nous avons par ailleurs observé une variabilité notable entre les différents microbiotes d'un même compartiment et d'un même génotype dans une même parcelle, c'est à dire à une échelle locale, voire à une échelle de l'individu. Une sélection exercée par l'hôte ne peut donc pas expliquer ce résultat. Ces variations de structure des communautés microbiennes non expliquées pourraient résulter de processus stochastiques liés à la dérive écologique et/ou à une limitation de la dispersion (Evans et al., 2017, Zhou et al., 2014).

Au niveau du sol des rizières, à petite échelle, les processus stochastiques tels qu'une réduction de la dispersion et une augmentation de la dérive jouent un rôle important dans la structuration

des communautés fongiques (Liu et al., 2020, Zhao et al., 2019). La combinaison de ces deux processus peut créer de la variation lors de la colonisation des niches générant une hétérogénéité lors de la succession due à des effets de priorité. Par ailleurs, les fréquentes inondations des rizières favorisent la connectivité hydrologique et l'atténuation des changements environnementaux subis par le sol et limitent *de facto* les possibles pressions de sélection (Liu et al., 2020, Zhao et al., 2019). Enfin, la présence de nombreux taxa rares ou peu abondants peut accentuer le phénomène de dérive (Zhou & Ning, 2017, Nemergut et al., 2013).

Il est ainsi fort probable que des processus stochastiques ont contribué à l'assemblage du microbiote du riz dans nos travaux. Nous n'avons cependant pas estimé dans ce travail de thèse l'importance des processus stochastiques dans l'assemblage des communautés microbiennes du riz. Pour inférer l'existence d'une stochasticité écologique dans l'assemblage de nos communautés, des approches statistiques à partir d'approches multivariées restent peu performantes pour indiquer l'influence et l'importance de la stochasticité. Ces approches mettent surtout en évidence l'importance relative de facteurs environnementaux en estimant l'importance de la sélection (Zhou & Ning, 2017). Une deuxième approche pour tester la présence de stochasticité, utilise des modèles de processus basés sur la théorie neutre qui cherchent à capturer l'influence de la diversification, de l'extinction, de la dispersion et de la dérive écologique sur l'abondance relative des espèces avec l'hypothèse que toutes les espèces à un niveau trophique particulier sont similaires au niveau démographique (Rosindell et al., 2011). Malgré parfois des hypothèses peu réalistes et une apparente simplicité, ces modèles neutres peuvent s'accorder avec de nombreux modèles écologiques (Zhou & Ning, 2017). Un point sensible à l'application de modèles neutres est l'effet de l'échantillonnage avec l'incertitude liée aux petites tailles d'échantillons dans une large population microbienne (Zhou & Ning, 2017). Une troisième approche serait l'utilisation de modèles nuls qui génèrent des profils stochastiques attendus statistiquement à partir de permutations aléatoires des données en excluant certains mécanismes. Quand les profils écologiques observés ne sont pas statistiquement différents des profils générés à partir du modèle nul, la dynamique de la communauté est considérée stochastique en respect des processus exclus (Zhou & Ning, 2017). L'analyse par des modèles nuls pose aussi certains challenges (Zhou & Ning, 2017).

Définir le rôle de la stochasticité et sa contribution dans la diversité des communautés microbiennes par des expérimentations au champ et au laboratoire est nécessaire. Installer deux parcelles homogènes en exerçant une pression de sélection ou en générant une perturbation sur l'une des deux pourrait peut-être permettre d'aborder cette question en contrôlant

l'hétérogénéité initiale et l'historique des parcelles. Des systèmes de microcosmes au laboratoire ou en serres pourraient aussi peut-être permettre de mieux gérer ces conditions environnementales avec des microbiotes synthétiques où certains systèmes pourraient recevoir des stress et d'autres non.

3.5.6. Vers une approche fonctionnelle du microbiote du riz ?

Contrairement à nos attentes, nos travaux n'ont pas mis en évidence une homogénéité taxonomique, à l'échelle de l'OTU, de la composition des communautés microbiennes d'échantillons d'un même génotype de riz, mais ont en revanche révélé une forte hétérogénéité de ces communautés microbiennes. De fait, il n'a pas été mis en évidence la présence notable d'un microbiote taxonomique cœur du riz à l'intérieur d'un génotype ou entre génotypes. Le microbiote taxonomique cœur du riz est ainsi limité à seulement quelques microorganismes, ce qui confirme des travaux antérieurs menés sur le riz (Edwards et al., 2015). Nous avons proposé dans le paragraphe précédent l'hypothèse d'assemblages stochastiques des communautés microbiennes pour expliquer leur forte variabilité et l'absence de microbiote taxonomique cœur.

Toutefois, l'importance des processus stochastiques en écologie pour expliquer la variabilité taxonomique observée reste vivement débattue (Zhou & Ning, 2017). On peut en effet, questionner la validité biologique de rechercher des traces de processus déterministes ou stochastiques dans l'assemblage des communautés microbiennes à partir de jeux de données majoritairement composées d'OTUs taxonomiques. Faire ce genre d'analyse est basé sur le postulat que la sélection des communautés microbiennes se fait sur la base de l'identité taxonomique, *a minima* à l'échelle spécifique ou plus sûrement à une échelle infraspécifique. Or, la taxonomie et/ou la phylogénie des microorganismes permettent leur identification taxonomique et/ou leur groupement phylogénétique en fonction de leurs séquences ADN, mais ne peuvent révéler leurs capacités à s'adapter à un environnement. Des microorganismes peuvent avoir une séquence nucléotidique du gène de l'ARNr 16S différente et partager plusieurs fonctions. Des bactéries différentes peuvent donc remplir les mêmes fonctions indispensables pour la plante, ce que l'on qualifie de redondance fonctionnelle.

A partir de ce type de critique, plusieurs groupes de recherche ont émis des doutes sur le postulat que la sélection des communautés microbiennes se ferait sur la base de l'identité taxonomique et ont proposé, sur la base de la théorie énoncée par Dawkins (« le gène égoïste », (Dawkin, 1976)) que la structuration des communautés microbiennes de la plante serait façonnée sur la

base de la structure fonctionnelle de celles-ci (Lemanceau et al., 2017). La théorie de Dawkins propose en effet que les organismes, ici les microbes, sont des véhicules de fonctions codées par des « réplicateurs » (les gènes). La plante attirerait ainsi des fonctions véhiculées par les microbes et dont elle a besoin par le jeu d'un dialogue moléculaire entre les molécules produites par elle-même et par ces mêmes microbes. Les gènes associés aux fonctions convoitées par la plante sont le plus souvent redondants dans un environnement donné et sont portés par des microbes taxonomiquement différents. La redondance des gènes entre les microorganismes pourrait expliquer l'hétérogénéité des communautés microbiennes associées à des individus d'un même génotype de l'hôte et à l'absence de microbiote taxonomique cœur.

L'analyse des profils fonctionnels des communautés microbiennes pourrait montrer que des processus déterministes sélectionnent les microorganismes en fonction de leurs capacités fonctionnelles (Tian et al., 2017). On pourrait s'attendre alors, dans le cas où les plantes sont proches phylogénétiquement, à ce qu'un microbiote cœur fonctionnel exhibe un plus grand degré de similarité. Le fait que certains groupes fonctionnels soient systématiquement associés à certains hôtes, suggère aussi que le microbiote cœur recruté est, d'une certaine façon, stable dans le temps (Lemanceau et al., 2017). Bien que nous ayons essayé dans nos travaux d'étudier les profils fonctionnels des communautés microbiennes associées au riz, l'état actuel des techniques et des bases de données fonctionnelles ne nous ont pas permis de confirmer ou d'infirmer la présence de structures fonctionnelles spécifiques parmi les communautés microbiennes identifiées pour chacun des génotypes de riz analysé. De futurs travaux devraient pourtant essayer d'approfondir cette piste de recherche comme cela a été fait dans plusieurs études pionnières récentes qui ont cherché à découpler les aspects taxonomiques et fonctionnels. A titre d'exemple, il a été montré que les plantes sélectionnent des fonctions portées par leurs microbiotes phyllosphériques (Bouffaud et al., 2016). Il a par ailleurs été révélé que le génotype de *Hordeum vulgare* avait peu d'effet structurant sur la composition taxonomique de son microbiote, mais qu'il impactait plus profondément les capacités fonctionnelles de son microbiote, et qu'une partie de cette diversification fonctionnelle était le résultat d'une sélection positive ou diversifiante (Bulgarelli et al., 2015). De même, des différences de microbiotes taxonomiques associés à différents échantillons de l'algue *Ulva australis* ne présentant que 15% de similitude taxonomique, ont conduit à l'hypothèse d'un modèle de loterie dans la colonisation de cette algue (Burke et al., 2011). En revanche, une forte similarité de la composition fonctionnelle a été observée (70%) et un microbiote cœur fonctionnel a été identifié présent chez toutes les communautés d'algues consistant avec leur

écologie (Burke et al., 2011). Dans une autre étude, des communautés bactériennes du rhizoplan de blé et du concombre ayant poussé dans un même sol, sont significativement différenciées au niveau taxonomique (Ofek-Lalzar et al., 2014). Cependant, l'analyse métagénomique a révélé un profil génétique fonctionnel similaire pour les microbiotes du rhizoplan des deux hôtes, et a identifié un ensemble de gènes fonctionnels associé à la colonisation des racines partagé entre les deux hôtes (Ofek-Lalzar et al., 2014).

Références bibliographiques

- Achouak W, Abrouk D, Guyonnet J, *et al.*, 2019. Plant hosts control microbial denitrification activity. *FEMS Microbiol Ecol* **95**.
- Agler MT, Ruhe J, Kroll S, *et al.*, 2016. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *PLoS Biol* **14**, e1002352.
- Ahmad Z, Sarwar M, Javed M, 2006. Agrep -- A Fast Approximate Pattern-Matching Tool.
- Aird D, Ross MG, Chen W-S, *et al.*, 2011. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biology* **12**, R18.
- Alcon-Giner C, Caim S, Mitra S, *et al.*, 2017. Optimisation of 16S rRNA gut microbiota profiling of extremely low birth weight infants. *BMC Genomics* **18**, 841.
- Allander T, Emerson SU, Engle RE, Purcell RH, Bukh J, 2001. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proceedings of the National Academy of Sciences* **98**, 11609-14.
- Almaguer M, Rojas TI, Rodríguez-Rajo FJ, Aira MJ, 2012. Airborne fungal succession in a rice field of Cuba. *European Journal of Plant Pathology* **133**, 473-82.
- Alonso P, Gladioux P, Moubset O, *et al.*, 2019. Emergence of Southern Rice Black-Streaked Dwarf Virus in the Centuries-Old Chinese Yuanyang Agrosystem of Rice Landraces. *Viruses* **11**, 985.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, 1990. Basic local alignment search tool. *J Mol Biol* **215**, 403-10.
- Amend AS, Seifert KA, Bruns TD, 2010. Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Molecular Ecology* **19**, 5555-65.
- Andreo-Jimenez B, Vandenkoornhuysen P, Lê Van A, *et al.*, 2019. Plant host and drought shape the root associated fungal microbiota in rice. *PeerJ* **7**, e7463-e.
- Angly FE, Dennis PG, Skarshewski A, Vanwonterghem I, Hugenholtz P, Tyson GW, 2014. CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome* **2**, 11.
- Ashkani S, Rafii MY, Shabanimofrad M, *et al.*, 2015. Molecular Breeding Strategy and Challenges Towards Improvement of Blast Disease Resistance in Rice Crop. *Front Plant Sci* **6**.
- Aulakh M, Wassmann R, Bueno C, Kreuzwieser J, Rennenberg H, 2001. Characterization of Root Exudates at Different Growth Stages of Ten Rice (*Oryza sativa* L.) Cultivars. *Plant Biology* **3**, 139-48.
- Aziz RK, 2009. A hundred-year-old insight into the gut microbiome! *Gut Pathog* **1**, 21-.
- Bachy C, Dolan JR, López-García P, Deschamps P, Moreira D, 2013. Accuracy of protist diversity assessments: morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. *The ISME journal* **7**, 244-55.
- Bacilio-Jiménez M, Aguilar-Flores S, Del Valle MV, Pérez A, Zepeda A, Zenteno E, 2001. Endophytic bacteria in rice seeds inhibit early colonization of roots by *Azospirillum brasilense*. *Soil Biology and Biochemistry* **33**, 167-72.
- Bacilio-Jiménez M, Aguilar-Flores S, Ventura-Zapata E, Pérez-Campos E, Bouquelet S, Zenteno E, 2003. Chemical characterization of root exudates from rice (*Oryza sativa*) and their effects on the chemotactic response of endophytic bacteria. *Plant and Soil* **249**, 271-7.
- Bacilio M, Aguilar-Flores S, Elsa V, Perez-Campos E, Bouquelet S, Zenteno E, 2003. Chemical characterization of root exudates from rice (*Oryza sativa*) and their effects on the chemotactic response of endophytic bacteria. *Plant and Soil* **249**, 271-7.

- Bae H-S, Rash BA, Rainey FA, *et al.*, 2007. Description of *Azospira restricta* sp. nov., a nitrogen-fixing bacterium isolated from groundwater. *International Journal of Systematic and Evolutionary Microbiology* **57**, 1521-6.
- Baedke J, Fábregas-Tejeda A, Nieves Delgado A, 2020. The holobiont concept before Margulis. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **n/a**.
- Bai Y, Muller DB, Srinivas G, *et al.*, 2015. Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* **528**, 364-9.
- Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM, 2006. THE ROLE OF ROOT EXUDATES IN RHIZOSPHERE INTERACTIONS WITH PLANTS AND OTHER ORGANISMS. *Annu Rev Plant Biol* **57**, 233-66.
- Baker TS, Olson NH, Fuller SD, 1999. Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiology and Molecular Biology Reviews* **63**, 862-922, table of contents.
- Bankevich A, Nurk S, Antipov D, *et al.*, 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of computational biology : a journal of computational molecular cell biology* **19**, 455-77.
- Bardenas EA, 1965. *Morphology and Varietal Characteristics of the Rice Plant, The*. Int. Rice Res. Inst.
- Bardgett RD, Van Der Putten WH, 2014. Belowground biodiversity and ecosystem functioning. *Nature* **515**, 505-11.
- Bashan Y, De-Bashan L, 2010. How the Plant Growth-Promoting Bacterium *Azospirillum* Promotes Plant Growth—A Critical Assessment. *Advances in Agronomy - ADVAN AGRON* **108**, 77-136.
- Bashan Y, Holguin G, De-Bashan LE, 2004. *Azospirillum*-plant relationships: physiological, molecular, agricultural, and environmental advances (1997-2003). *Can J Microbiol* **50**, 521-77.
- Bass D, Stentiford GD, Wang H-C, Koskella B, Tyler CR, 2019. The Pathobiome in Animal and Plant Diseases. *Trends in Ecology and Evolution* **34**, 996-1008.
- Bastias DA, Johnson LJ, Card SD, 2019. Symbiotic bacteria of plant-associated fungi: friends or foes? *Curr Opin Plant Biol* **56**, 1-8.
- Bellemain E, Carlsen T, Brochmann C, Coissac E, Taberlet P, Kauserud H, 2010. ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiol* **10**, 189.
- Berendsen RL, Pieterse CM, Bakker PA, 2012. The rhizosphere microbiome and plant health. *Trends Plant Sci* **17**, 478-86.
- Berg G, 2009. Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. *Appl Microbiol Biotechnol* **84**, 11-8.
- Berg G, Grube M, Schlöter M, Smalla K, 2014a. The plant microbiome and its importance for plant and human health. *Front Microbiol* **5**.
- Berg G, Grube M, Schlöter M, Smalla K, 2014b. Unraveling the plant microbiome: looking back and future perspectives. *Front Microbiol* **5**.
- Berg G, Raaijmakers JM, 2018. Saving seed microbiomes. *The ISME journal* **12**, 1167-70.
- Berg G, Smalla K, 2009. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol Ecol* **68**, 1-13.
- Berg M, Koskella B, 2018. Nutrient- and Dose-Dependent Microbiome-Mediated Protection against a Plant Pathogen. *Curr Biol* **28**, 2487-92.e3.

- Bernardo P, Charles-Dominique T, Barakat M, *et al.*, 2018. Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *The ISME journal* **12**, 173-84.
- Bernatzky R, Tanksley SD, 1986. TOWARD A SATURATED LINKAGE MAP IN TOMATO BASED ON ISOZYMES AND RANDOM cDNA SEQUENCES. *Genetics* **112**, 887-98.
- Bertani I, Abbruscato P, Piffanelli P, Subramoni S, Venturi V, 2016. Rice bacterial endophytes: isolation of a collection, identification of beneficial strains and microbiome analysis. *Environ Microbiol Rep* **8**, 388-98.
- Bilen M, Dufour J-C, Lagier J-C, *et al.*, 2018. The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species. *Microbiome* **6**, 94.
- Biscarini F, Cozzi P, Casella L, *et al.*, 2016. Genome-Wide Association Study for Traits Related to Plant and Grain Morphology, and Root Architecture in Temperate Rice Accessions. *PLoS One* **11**, e0155425.
- Bolduc J-S, 2012. Traduire la notion de fitness en français. *Bulletin d'histoire et d'épistémologie des sciences de la vie* **Volume 19**, 67-96.
- Bonder MJ, Abeln S, Zaura E, Brandt BW, 2012. Comparing clustering and pre-processing in taxonomy analysis. *Bioinformatics* **28**, 2891-7.
- Borbet TC, Blaser MJ, 2019. Host genotype and early life microbiota alterations have additive effects on disease susceptibility. *Mucosal Immunology* **12**, 586-8.
- Bos L, 1977. Persistence of infectivity of three viruses in plant material dried over CaCl₂ and stored under different conditions. *Netherlands Journal of Plant Pathology* **83**, 217-20.
- Bouffaud ML, Poirier MA, Muller D, Moenne-Loccoz Y, 2014. Root microbiome relates to plant host evolution in maize and other Poaceae. *Environ Microbiol* **16**, 2804-14.
- Bouffaud ML, Renoud S, Moenne-Loccoz Y, Muller D, 2016. Is plant evolutionary history impacting recruitment of diazotrophs and nifH expression in the rhizosphere? *Sci Rep* **6**, 21690.
- Brader G, Compant S, Vescio K, *et al.*, 2017. Ecology and genomic insights into plant-pathogenic and plant-nonpathogenic endophytes. *Annual Review of Phytopathology* **55**, 61-83.
- Bray JR, Curtis JT, 1957. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* **27**, 325-49.
- Breidenbach B, Pump J, Dumont MG, 2016. Microbial Community Structure in the Rhizosphere of Rice Plants. *Front Microbiol* **6**, 1537-.
- Breitbart M, Rohwer F, 2005. Here a virus, there a virus, everywhere the same virus? *Trends in Microbiology* **13**, 278-84.
- Breitbart M, Salamon P, Andresen B, *et al.*, 2002. Genomic analysis of uncultured marine viral communities. *Proceedings of the National Academy of Sciences* **99**, 14250-5.
- Brister JR, Ako-Adjei D, Bao Y, Blinkova O, 2015. NCBI viral genomes resource. *Nucleic Acids Res* **43**, D571-7.
- Brooks JP, Edwards DJ, Harwich MD, *et al.*, 2015. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol* **15**, 66.
- Bulgarelli D, Garrido-Oter R, Münch PC, *et al.*, 2015. Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host & Microbe* **17**, 392-403.
- Bulgarelli D, Rott M, Schlaeppi K, *et al.*, 2012. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* **488**, 91-5.

- Bulgarelli D, Schlaeppi K, Spaepen S, Van Themaat EVL, Schulze-Lefert P, 2013. Structure and Functions of the Bacterial Microbiota of Plants. *Annu Rev Plant Biol* **64**, 807-38.
- Burke C, Steinberg P, Rusch D, Kjelleberg S, Thomas T, 2011. Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences* **108**, 14288-93.
- Busby PE, Ridout M, Newcombe G, 2016. Fungal endophytes: modifiers of plant disease. *Plant Molecular Biology* **90**, 645-55.
- Buttimer C, McAuliffe O, Ross RP, Hill C, O'mahony J, Coffey A, 2017. Bacteriophages and Bacterial Plant Diseases. *Front Microbiol* **8**.
- Callahan BJ, McMurdie PJ, Holmes SP, 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal* **11**, 2639-43.
- Caporaso JG, Kuczynski J, Stombaugh J, *et al.*, 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335-6.
- Case RJ, Boucher Y, Dahllöf I, Holmström C, Doolittle WF, Kjelleberg S, 2007. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Applied and Environmental Microbiology* **73**, 278-88.
- Castrillo G, Teixeira PJPL, Paredes SH, *et al.*, 2017. Root microbiota drive direct integration of phosphate stress and immunity. *Nature* **543**, 513-8.
- Chaibub AA, Sousa TPD, Araújo LGD, Filippi MCCD, 2020. Molecular and morphological characterization of rice phylloplane fungi and determination of the antagonistic activity against rice pathogens. *Microbiol Res* **231**, 126353.
- Chaparro JM, Badri DV, Vivanco JM, 2014. Rhizosphere microbiome assemblage is affected by plant development. *ISME J* **8**, 790-803.
- Chi F, Shen S-H, Cheng H-P, Jing Y-X, Yanni YG, Dazzo FB, 2005. Ascending migration of endophytic rhizobia, from roots to leaves, inside rice plants and assessment of benefits to rice growth physiology. *Applied and Environmental Microbiology* **71**, 7271-8.
- Chinnadurai C, Balachandar D, Sundaram SP, 2009. Characterization of 1-aminocyclopropane-1-carboxylate deaminase producing methylobacteria from phyllosphere of rice and their role in ethylene regulation. *World Journal of Microbiology and Biotechnology* **25**, 1403-11.
- Ciancio DA, Mukerji KG, 2008. *Integrated Management of Diseases Caused by Fungi, Phytoplasma and Bacteria*.
- Clay K, Schardl C, 2002. Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. *The American Naturalist* **160**, S99-S127.
- Coleman-Derr D, Desgarennes D, Fonseca-Garcia C, *et al.*, 2016. Plant compartment and biogeography affect microbiome composition in cultivated and native Agave species. *New Phytol* **209**, 798-811.
- Compant S, Clément C, Sessitsch A, 2010. Plant Growth-Promoting Bacteria in the Rhizo- and Endosphere of Plants: Their Role, Colonization, Mechanisms Involved and Prospects for Utilization. *Soil Biology and Biochemistry*, 669-78.
- Compant S, Mitter B, Colli-Mull JG, Gangl H, Sessitsch A, 2011. Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches of colonization. *Microb Ecol* **62**, 188-97.
- Compant S, Samad A, Faist H, Sessitsch A, 2019. A review on the plant microbiome: Ecology, functions and emerging trends in microbial application. *Journal of Advanced Research* **19**.
- Copeland JK, Yuan L, Layeghifard M, Wang PW, Guttman DS, 2015. Seasonal community succession of the phyllosphere microbiome. *Mol Plant Microbe Interact* **28**, 274-85.

- Cordovez V, Dini-Andreote F, Carrión VJ, Raaijmakers JM, 2019. Ecology and Evolution of Plant Microbiomes. *Annual Review of Microbiology* **73**, 69-88.
- Cottyn B, Debode J, Regalado E, Mew TW, Swings J, 2009. Phenotypic and genetic diversity of rice seed-associated bacteria and their role in pathogenicity and biological control. *Journal of Applied Microbiology* **107**, 885-97.
- Cottyn B, Regalado E, Lanoot B, De Cleene M, Mew TW, Swings J, 2001. Bacterial populations associated with rice seed in the tropical environment. *Phytopathology* **91**, 282-92.
- Cregger MA, Veach AM, Yang ZK, *et al.*, 2018. The Populus holobiont: dissecting the effects of plant niches and genotype on the microbiome. *Microbiome* **6**, 31.
- Dawkin R, 1976. The selfish gene. *Oxford University Press* **1**, 976.
- De La Paz Giménez-Pecchi M, Bogo MR, Santi L, *et al.*, 2002. Characterization of Mycoviruses and Analyses of Chitinase Secretion in the Biocontrol Fungus *Metarhizium anisopliae*. *Curr Microbiol* **45**, 334-9.
- Dedeurwaerdere T, Hannachi M, 2019. Socio-economic drivers of coexistence of landraces and modern crop varieties in agro-biodiversity rich Yunnan rice fields. *Ecological Economics* **159**, 177-88.
- Desmarais E, Lanneluc I, Lagnel J, 1998. Direct amplification of length polymorphisms (DALP), or how to get and characterize new genetic markers in many species. *Nucleic Acids Research* **26**, 1458-65.
- Ding LJ, Cui HL, Nie SA, Long XE, Duan GL, Zhu YG, 2019. Microbiomes inhabiting rice roots and rhizosphere. *FEMS Microbiol Ecol* **95**.
- Dini-Andreote F, Stegen JC, Van Elsas JD, Salles JF, 2015. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences* **112**, E1326-E32.
- Dombrowski N, Schlaeppi K, Agler MT, *et al.*, 2017. Root microbiota dynamics of perennial *Arabidopsis alpina* are dependent on soil residence time but independent of flowering time. *ISME J.* **11**, 43-55. doi: 10.1038/ismej.2016.109. Epub Aug 2.
- Drogue B, Sanguin H, Chamam A, *et al.*, 2014. Plant root transcriptome profiling reveals a strain-dependent response during *Azospirillum*-rice cooperation. *Front Plant Sci* **5**, 607.
- Du H, Yu Y, Ma Y, *et al.*, 2017. Sequencing and de novo assembly of a near complete indica rice genome. *Nat Commun* **8**, 15324.
- Edgar RC, 2018. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* **34**, 2371-5.
- Edwards J, Johnson C, Santos-Medellin C, *et al.*, 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci U S A* **112**, E911-20.
- Edwards J, Santos-Medellín C, Nguyen B, *et al.*, 2019. Soil domestication by rice cultivation results in plant-soil feedback through shifts in soil microbiota. *Genome Biology* **20**, 221.
- Edwards JA, Santos-Medellin CM, Liechty ZS, *et al.*, 2018. Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. *PLoS Biology* **16**, e2003862. doi: 10.1371/journal.pbio.. eCollection 2018 Feb.
- Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS, 2019. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends in Microbiology* **27**, 105-17.
- Elshire RJ, Glaubitz JC, Sun Q, *et al.*, 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**, e19379.

- Emerson JB, Adams RI, Román CMB, *et al.*, 2017. Schrödinger's microbes: Tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome* **5**, 86.
- Engel P, Moran NA, 2013. The gut microbiota of insects – diversity in structure and function. *FEMS Microbiol Rev* **37**, 699-735.
- Eren AM, Maignien L, Sul WJ, *et al.*, 2013. Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods in ecology and evolution* **4**, 1111-9.
- Escudié F, Auer L, Bernard M, *et al.*, 2017. FROGS: Find, Rapidly, OTUs with Galaxy Solution. *Bioinformatics* **34**, 1287-94.
- Etesami H, Alikhani H, 2016. Co-inoculation with endophytic and rhizosphere bacteria allows reduced application rates of N-fertilizer for rice plant (*Oryza sativa* L.). *Rhizosphere* **2**.
- Evans S, Martiny JBH, Allison SD, 2017. Effects of dispersal and selection on stochastic assembly in microbial communities. *The ISME journal* **11**, 176-85.
- Eyre A, Wang M, Oh Y, Dean R, 2019. Identification and Characterization of the Core Rice Seed Microbiome. *Phytobiomes Journal* **3**.
- Fancello L, Raoult D, Desnues C, 2012. Computational tools for viral metagenomics and their application in clinical research. *Virology* **434**, 162-74.
- Farré-Maduell E, Casals-Pascual C, 2019. The origins of gut microbiome research in Europe: From Escherich to Nissle. *Human Microbiome Journal* **14**, 100065.
- Feng Y, Shen D, Dong X, Song W, 2003. In vitro symplasmata formation in the rice diazotrophic endophyte *Pantoea agglomerans* YS19. *Plant and Soil* **255**, 435-44.
- Fierer N, Nemergut D, Knight R, Craine JM, 2010. Changes through time: integrating microorganisms into the study of succession. *Res Microbiol* **161**, 635-42.
- Filloux D, Dallot S, Delaunay A, Galzi S, Jacquot E, Roumagnac P, 2015. Metagenomics Approaches Based on Virion-Associated Nucleic Acids (VANA): An Innovative Tool for Assessing Without A Priori Viral Diversity of Plants. *Methods Mol Biol* **1302**:249-57., 10.1007/978-1-4939-2620-6_18.
- Fitzpatrick CR, Lu-Irving P, Copeland J, *et al.*, 2018. Chloroplast sequence variation and the efficacy of peptide nucleic acids for blocking host amplification in plant microbiome studies. *Microbiome* **6**, 144.
- Forterre P, 2016. To be or not to be alive: How recent discoveries challenge the traditional definitions of viruses and life. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* **59**, 100-8.
- Foster ZS, Sharpton TJ, Grunwald NJ, 2017. Metacoder: An R package for visualization and manipulation of community taxonomic diversity data. *PLoS Comput Biol* **13**, e1005404.
- Fox GE, Magrum LJ, Balch WE, Wolfe RS, Woese CR, 1977. Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *Proceedings of the National Academy of Sciences* **74**, 4537-41.
- François S, Filloux D, Fernandez E, Ogliastro M, Roumagnac P, 2018. Viral Metagenomics Approaches for High-Resolution Screening of Multiplexed Arthropod and Plant Viral Communities. In., 77-95. (1746.)
- Frank AC, Saldierna Guzmán JP, Shay JE, 2017. Transmission of bacterial endophytes. *Microorganisms* **5**, 70.
- Fukuhara T, 2019. Endornaviruses: persistent dsRNA viruses with symbiotic properties in diverse eukaryotes. *Virus Genes* **55**, 165-73.
- Fukuhara T, Koga R, Aoki N, *et al.*, 2006. The wide distribution of endornaviruses, large double-stranded RNA replicons with plasmid-like properties. *Archives of Virology* **151**, 995-1002.

- Ganeshaiah K, Sanjappa M, Rao R, Murugan C, Shivaprakash K, 2019. Spatial distribution pattern of taxonomic and phylogenetic diversity of woody flora in Andaman and Nicobar Islands, India. *Forest Ecosystems* **6**, 38.
- Garbaye J, 1994. Helper bacteria: a new dimension to the mycorrhizal symbiosis. *New Phytol* **128**, 197-210.
- Gilbert SF, 2019. Evolutionary transitions revisited: Holobiont evo-devo. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **332**, 307-14.
- Gitaitis RD, Walcott RR, 2007. The epidemiology and management of seedborne bacterial diseases. *Annual Review of Phytopathology* **45**, 371-97.
- Gopal M, Gupta A, 2016. Microbiome Selection Could Spur Next-Generation Plant Breeding Strategies. *Front Microbiol* **7**.
- Gorovits R, Sobol I, Altaieb M, Czosnek H, Anfoka G, 2019. Taking advantage of a pathogen: understanding how a virus alleviates plant stress response. *Phytopathology Research* **1**, 20.
- Gouba N, Raoult D, Drancourt M, 2013. Plant and fungal diversity in gut microbiota as revealed by molecular and culture investigations. *PLoS One* **8**, e59474-e.
- Graham EB, Knelman JE, Schindlbacher A, *et al.*, 2016. Microbes as Engines of Ecosystem Function: When Does Community Structure Enhance Predictions of Ecosystem Processes? *Front Microbiol* **7**.
- Graspeuntner S, Loeper N, Künzel S, Baines JF, Rupp J, 2018. Selection of validated hypervariable regions is crucial in 16S-based microbiota studies of the female genital tract. *Sci Rep* **8**, 9678.
- Green SJ, Inbar E, Michel FC, Jr., Hadar Y, Minz D, 2006. Succession of bacterial communities during early plant development: transition from seed to root and effect of compost amendment. *Applied and Environmental Microbiology* **72**, 3975-83.
- Guerrero R, Margulis L, Berlanga M, 2013. Symbiogenesis: the holobiont as a unit of evolution. *Int Microbiol* **16**, 133-43.
- Gury J, Zinger L, Gielly L, Taberlet P, Geremia RA, 2008. Exonuclease activity of proofreading DNA polymerases is at the origin of artifacts in molecular profiling studies. *Electrophoresis* **29**, 2437-44.
- Hacquard S, Garrido-Oter R, González A, *et al.*, 2015. Microbiota and Host Nutrition across Plant and Animal Kingdoms. *Cell Host & Microbe* **17**, 603-16.
- Halary S, Temmam S, Raoult D, Desnues C, 2016. Viral metagenomics: are we missing the giants? *Curr Opin Microbiol* **31**, 34-43.
- Hallmann J, Quadat A, Mahaffee W, Kloepper J, 2011. Endophytic Bacteria in Agricultural Crops. *Can J Microbiol* **43**, 895-914.
- Hamonts K, Trivedi P, Garg A, *et al.*, 2018. Field study reveals core plant microbiota and relative importance of their drivers. *Environ Microbiol* **20**, 124-40.
- Handelsman J, Rondon M, Brady S, Clardy J, Goodman R, 1998. Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM.. Molecular Biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* **5**: R245-R249. *Chemistry & biology* **5**, R245-9.
- Haney C, Samuel B, Bush J, Ausubel F, 2015. Associations with rhizosphere bacteria can confer an adaptive advantage to plants. *Nature Plants* **1**, 15051.
- Hardoim PR, Andreote FD, Reinhold-Hurek B, Sessitsch A, Van Overbeek LS, Van Elsas JD, 2011. Rice root-associated bacteria: insights into community structures across 10 cultivars. *FEMS Microbiol Ecol* **77**, 154-64.

- Hardoim PR, Hardoim CC, Van Overbeek LS, Van Elsas JD, 2012. Dynamics of seed-borne rice endophytes on early plant growth stages. *PLoS One* **7**, e30438. doi: 10.1371/journal.pone.0030438. Epub 2012 Feb 17.
- Hardoim PR, Van Overbeek LS, Berg G, *et al.*, 2015. The Hidden World within Plants: Ecological and Evolutionary Considerations for Defining Functioning of Microbial Endophytes. *Microbiology and Molecular Biology Reviews* **79**, 293-320.
- Hardoim PR, Van Overbeek LS, Elsas JDV, 2008. Properties of bacterial endophytes and their proposed role in plant growth. *Trends in Microbiology* **16**, 463-71.
- Hartman K, Van Der Heijden MGA, Wittwer RA, Banerjee S, Walser JC, Schlaeppi K, 2018. Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming (vol 6, 14, 2018). *Microbiome* **6**.
- Hassani MA, Durán P, Hacquard S, 2018. Microbial interactions within the plant holobiont. *Microbiome* **6**, 58.
- Hayes S, Mahony J, Nauta A, Van Sinderen D, 2017. Metagenomic Approaches to Assess Bacteriophages in Various Environmental Niches. *Viruses* **9**.
- He J, Zhao X, Laroche A, Lu ZX, Liu H, Li Z, 2014. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front Plant Sci* **5**, 484.
- Hoang AT, Zhang HM, Yang J, *et al.*, 2011. Identification, Characterization, and Distribution of Southern rice black-streaked dwarf virus in Vietnam. *Plant Disease* **95**, 1063-9.
- Holguin G, Patten C, Glick B, 1999. Genetics and molecular biology of Azospirillum. *Biology and Fertility of Soils* **29**.
- Honjo MN, Emura N, Kawagoe T, *et al.*, 2020. Seasonality of interactions between a plant virus and its host during persistent infection in a natural environment. *The ISME journal* **14**, 506-18.
- Hooper DU, Chapin Iii FS, Ewel JJ, *et al.*, 2005. EFFECTS OF BIODIVERSITY ON ECOSYSTEM FUNCTIONING: A CONSENSUS OF CURRENT KNOWLEDGE. *Ecological Monographs* **75**, 3-35.
- Horiuchi H, Moriyama H, Fukuhara T, 2003. Inheritance of Oryza sativa endornavirus in F1 and F2 hybrids between japonica and indica rice. *Genes Genet Syst* **78**, 229-34.
- Huang X, Madan A, 1999. CAP3: A DNA sequence assembly program. *Genome Res* **9**, 868-77.
- Hussain Q, Liu Y, Zhang A, *et al.*, 2011. Variation of bacterial and fungal community structures in the rhizosphere of hybrid and standard rice cultivars and linkage to CO2 flux. *FEMS Microbiol Ecol* **78**, 116-28.
- Hussain Q, Pan G, Liu YZ, *et al.*, 2012. Microbial community dynamics and function associated with rhizosphere over periods of rice growth. *Plant Soil and Environment* **58**, 55-61.
- Imchen M, Kumavath R, Vaz ABM, *et al.*, 2019. 16S rRNA Gene Amplicon Based Metagenomic Signatures of Rhizobiome Community in Rice Field During Various Growth Stages. *Front Microbiol* **10**.
- Iwen PC, Hinrichs SH, Rupp ME, 2002. Utilization of the internal transcribed spacer regions as molecular targets to detect and identify human fungal pathogens. *Medical Mycology* **40**, 87-109.
- Jaccard P, 1912. THE DISTRIBUTION OF THE FLORA IN THE ALPINE ZONE.1. *New Phytologist* **11**, 37-50.
- Jha PN, Gomaa AB, Yanni YG, *et al.*, 2019. Alterations in the Endophyte-Enriched Root-Associated Microbiome of Rice Receiving Growth-Promoting Treatments of Urea Fertilizer and Rhizobium Biofertilizer. *Microb Ecol*.
- Jiao Y, Li X, Liang L, *et al.*, 2012. Indigenous ecological knowledge and natural resource management in the cultural landscape of China's Hani Terraces. *Ecological Research* **27**, 247-63.

- Johnson JS, Spakowicz DJ, Hong B-Y, *et al.*, 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun* **10**, 5029.
- Jones MS, Kapoor A, Lukashov VV, Simmonds P, Hecht F, Delwart E, 2005. New DNA Viruses Identified in Patients with Acute Viral Infection Syndrome. *Journal of Virology* **79**, 8230-6.
- Kaga H, Mano H, Tanaka F, Watanabe A, Kaneko S, Morisaki H, 2009. Rice Seeds as Sources of Endophytic Bacteria. *Microbes and Environments* **24**, 154-62.
- Katsantonis D, Kadoglidou K, Dramalis C, Puigdollers, 2017. Rice blast forecasting models and their practical value: A review. *Phytopathologia Mediterranea* **56**, 187-216.
- Khan MU, Sessitsch A, Harris M, *et al.*, 2015. Cr-resistant rhizo-and endophytic bacteria associated with *Prosopis juliflora* and their potential as phytoremediation enhancing agents in metal-degraded soils. *Front Plant Sci* **5**, 755.
- Khankhum S, Valverde RA, 2018. Physiological traits of endornavirus-infected and endornavirus-free common bean (*Phaseolus vulgaris*) cv Black Turtle Soup. *Arch Virol* **163**, 1051-6.
- Kim H, Lee Y-H, 2019. The Rice Microbiome: A Model Platform for Crop Holobiome. *Phytobiomes Journal*.
- Kim K-H, Bae J-W, 2011. Amplification Methods Bias Metagenomic Libraries of Uncultured Single-Stranded and Double-Stranded DNA Viruses. *Applied and Environmental Microbiology* **77**, 7663-8.
- Knauth S, Hurek T, Brar D, Reinhold-Hurek B, 2005. Influence of different *Oryza* cultivars on expression of *nifH* gene pools in roots of rice. *Environ Microbiol* **7**, 1725-33.
- Knief C, Delmotte N, Chaffron S, *et al.*, 2012. Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *The ISME journal* **6**, 1378-90.
- Knudsen B, Bergmark L, Munk P, *et al.*, 2016. Impact of Sample Type and DNA Isolation Procedure on Genomic Inference of Microbiome Composition. *mSystems* **1**, e00095-16.
- Koskella B, Brockhurst MA, 2014. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol Rev.* **38**, 916-31. doi: 10.1111/574-6976.12072. Epub 2014 Mar 27.
- Kreznar JH, Keller MP, Traeger LL, *et al.*, 2017. Host Genotype and Gut Microbiome Modulate Insulin Secretion and Diet-Induced Metabolic Phenotypes. *Cell Reports* **18**, 1739-50.
- Krishnamurthy S, Wang D, 2017. Origins and Challenges of Viral Dark Matter. *Virus Research* **239**.
- Kurokawa Y, Nagai K, Huan PD, *et al.*, 2018. Rice leaf hydrophobicity and gas films are conferred by a wax synthesis gene (LGF1) and contribute to flood tolerance. *New Phytologist* **218**, 1558-69.
- Lagier J-C, Armougom F, Million M, *et al.*, 2012. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clinical Microbiology and Infection* **18**, 1185-93.
- Lagier J-C, Dubourg G, Million M, *et al.*, 2018. Culturing the human microbiota and culturomics. *Nature Reviews Microbiology* **16**, 540-50.
- Lagier J-C, Edouard S, Pagnier I, Mediannikov O, Drancourt M, Raoult D, 2015. Current and Past Strategies for Bacterial Culture in Clinical Microbiology. *Clinical Microbiology Reviews* **28**, 208-36.
- Lagier J-C, Khelaifia S, Alou MT, *et al.*, 2016. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol* **1**, 16203.
- Lagkouvardos I, Overmann J, Clavel T, 2017. Cultured microbes represent a substantial fraction of the human and mouse gut microbiota. *Gut Microbes* **8**, 493-503.
- Lakshmanan V, Cottone J, Bais HP, 2016. Killing Two Birds with One Stone: Natural Rice Rhizospheric Microbes Reduce Arsenic Uptake and Blast Infections in Rice. *Front Plant Sci* **7**, 1514-.

- Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR, 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences* **82**, 6955-9.
- Langmead B, Salzberg SL, 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-9.
- Layeghifard M, Hwang DM, Guttman DS, 2017. Disentangling Interactions in the Microbiome: A Network Perspective. *Trends in Microbiology* **25**, 217-28.
- Lee HJ, Jeong SE, Kim PJ, Madsen EL, Jeon CO, 2015. High resolution depth distribution of Bacteria, Archaea, methanotrophs, and methanogens in the bulk and rhizosphere soils of a flooded rice paddy. *Front Microbiol* **6**, 639.
- Lemanceau P, Blouin M, Muller D, Moenne-Loccoz Y, 2017. Let the Core Microbiota Be Functional. *Trends Plant Sci* **22**, 583-95.
- Li J-Y, Wang J, Zeigler RS, 2014. The 3,000 rice genomes project: new opportunities and challenges for future rice research. *Gigascience* **3**, 2047-217X-3-8.
- Li YH, Zhu JN, Zhai ZH, Zhang Q, 2010. Endophytic bacterial diversity in roots of *Phragmites australis* in constructed Beijing Cuihu Wetland (China). *FEMS Microbiology Letters* **309**, 84-93.
- Liao J, Huang H, Meusnier I, *et al.*, 2016. Pathogen effectors and plant immunity determine specialization of the blast fungus to rice subspecies. *Elife* **5**.
- Liesack W, Schnell S, Revsbech NP, 2000. Microbiology of flooded rice paddies. *FEMS Microbiol Rev* **24**, 625-45.
- Lima MT, Andrade ACDSP, Oliveira GP, *et al.*, 2019. Virus and microbiota relationships in humans and other mammals: An evolutionary view. *Human Microbiome Journal* **11**, 100050.
- Lindow SE, Brandl MT, 2003. Microbiology of the phyllosphere. *Applied and Environmental Microbiology* **69**, 1875-83.
- Litt M, Luty JA, 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American journal of human genetics* **44**, 397-401.
- Liu J, Abdelfattah A, Norelli J, *et al.*, 2018. Apple endophytic microbiota of different rootstock/scion combinations suggests a genotype-specific influence. *Microbiome* **6**, 18.
- Liu W, Graham EB, Zhong L, *et al.*, 2020. Dynamic microbial assembly processes correspond to soil fertility in sustainable paddy agroecosystems. *Functional Ecology* **n/a**.
- Long HH, Sonntag DG, Schmidt DD, Baldwin IT, 2010. The structure of the culturable root bacterial endophyte community of *Nicotiana attenuata* is organized by soil composition and host plant ethylene production and perception. *New Phytologist* **185**, 554-67.
- Louca S, Doebeli M, Parfrey LW, 2018. Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome* **6**, 41-.
- Lozupone C, Hamady M, Kelley S, Knight R, 2007. Quantitative and Qualitative Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. *Applied and Environmental Microbiology* **73**, 1576-85.
- Lozupone C, Knight R, 2005. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology* **71**, 8228-35.
- Lu Y, Watanabe A, Kimura M, 2002. Contribution of plant-derived carbon to soil microbial biomass dynamics in a paddy rice microcosm. *Biology and Fertility of Soils* **36**, 136-42.

- Luijten MLGC, Roelofsen W, Langenhoff AaM, Schraa G, Stams AJM, 2004. Hydrogen threshold concentrations in pure cultures of halo-respiring bacteria and at a site polluted with chlorinated ethenes. *Environ Microbiol* **6**, 646-50.
- Lundberg DS, Lebeis SL, Paredes SH, *et al.*, 2012. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**, 86-90.
- Lundberg DS, Yourstone S, Mieczkowski P, Jones CD, Dangl JL, 2013. Practical innovations for high-throughput amplicon sequencing. *Nat Methods* **10**, 999-1002.
- Lv MF, Xie L, Wang HF, Wang HD, Chen JP, Zhang HM, 2017. Biology of Southern rice black-streaked dwarf virus: a novel fijivirus emerging in East Asia. *Plant Pathology* **66**, 515-21.
- Ma B, Lv X, Warren A, Gong J, 2013. Shifts in diversity and community structure of endophytic bacteria and archaea across root, stem and leaf tissues in the common reed, *Phragmites australis*, along a salinity gradient in a marine tidal wetland of northern China. *Antonie Van Leeuwenhoek* **104**, 759-68.
- Ma M, Du H, Sun T, An S, Yang G, Wang D, 2019. Characteristics of archaea and bacteria in rice rhizosphere along a mercury gradient. *Science of The Total Environment* **650**, 1640-51.
- Madhaiyan M, Poonguzhali S, Kwon S-W, Sa T-M, 2009. *Methylobacterium phyllosphaerae* sp. nov., a pink-pigmented, facultative methylotroph from the phyllosphere of rice. *International Journal of Systematic and Evolutionary Microbiology* **59**, 22-7.
- Madhaiyan M, Poonguzhali S, Murugaiyan S, *et al.*, 2004. Growth promotion and induction of systemic resistance in rice cultivar Co-47 (*Oryza sativa* L.) by *Methylobacterium* spp. *Botanical Bulletin-Academia Sinica Taipei* **45**, 315-24.
- Madhusoodanan J, 2019. News Feature: Do hosts and their microbes evolve as a unit? *Proceedings of the National Academy of Sciences* **116**, 14391-4.
- Mahe F, Rognes T, Quince C, De Vargas C, Dunthorn M, 2015. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* **3**, e1420.
- Maignien L, Deforce EA, Chafee ME, Eren AM, Simmons SL, 2014. Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* phyllosphere communities. *Mbio* **5**, e00682-13.
- Manandhar HK, Jorgensen HJL, Smedegaard-Petersen V, Mathur SB, 1998. Seedborne Infection of Rice by *Pyricularia oryzae* and Its Transmission to Seedlings. *Plant Disease* **82**, 1093-9.
- Mano H, Morisaki H, 2008. Endophytic Bacteria in the Rice Plant. *Microbes and Environments* **23**, 109-17.
- Mareque C, Da Silva TF, Vollu RE, Beracochea M, Seldin L, Battistoni F, 2018. The Endophytic Bacterial Microbiota Associated with Sweet Sorghum (*Sorghum bicolor*) Is Modulated by the Application of Chemical N Fertilizer to the Field. *International Journal of Genomics*.
- Margulis L, Fester R, 1991. Bellagio conference and book. Symbiosis as Source of Evolutionary Innovation: Speciation and Morphogenesis. Conference--June 25-30, 1989, Bellagio Conference Center, Italy. *Symbiosis (Philadelphia, Pa.)* **11**, 93-101.
- Martin F, Uroz S, Barker D, 2017. Ancestral alliances: Plant mutualistic symbioses with fungi and bacteria. *Science* **356**, eaad4501.
- Martin M, 2011. CUTADAPT removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**.
- Mbodj D, Effa-Effa B, Kane A, *et al.*, 2018. Arbuscular mycorrhizal symbiosis in rice: Establishment, environmental control and impact on plant growth and resistance to abiotic stresses. *Rhizosphere* **8**, 12-26.

- Mcmurdie PJ, Holmes S, 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217.
- Mendes R, Garbeva P, Raaijmakers JM, 2013. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol. Rev.* **37**, 634-63.
- Mendes R, Kruijt M, De Bruijn I, *et al.*, 2011. Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **332**, 1097.
- Mertens P, Maan S, Samuel A, *et al.*, 2005. Virus taxonomy: eighth report of the International Committee on Taxonomy of Viruses. *Fauquet, CM*, 447-54.
- Moriyama H, Horiuchi H, Koga R, Fukuhara T, 1999a. Molecular characterization of two endogenous double-stranded RNAs in rice and their inheritance by interspecific hybrids. *J Biol Chem* **274**, 6882-8.
- Moriyama H, Horiuchi H, Nitta T, Fukuhara T, 1999b. Unusual inheritance of evolutionarily-related double-stranded RNAs in interspecific hybrid between rice plants *Oryza sativa* and *Oryza rufipogon*. *Plant Mol Biol* **39**, 1127-36.
- Moriyama H, Kanaya K, Wang JZ, Nitta T, Fukuhara T, 1996. Stringently and developmentally regulated levels of a cytoplasmic double-stranded RNA and its high-efficiency transmission via egg and pollen in rice. *Plant Mol Biol* **31**, 713-9.
- Moriyama H, Urayama S-I, Higashiura T, Le TM, Komatsu K, 2018. Chrysovirus in *Magnaporthe oryzae*. *Viruses* **10**, 697.
- Moronta-Barrios F, Gionechetti F, Pallavicini A, Marys E, Venturi V, 2018. Bacterial Microbiota of Rice Roots: 16S-Based Taxonomic Profiling of Endophytic and Rhizospheric Diversity, Endophytes Isolation and Simplified Endophytic Community. *Microorganisms* **6**, 14.
- Moya A, Ferrer M, 2016. Functional Redundancy-Induced Stability of Gut Microbiota Subjected to Disturbance. *Trends in Microbiology* **24**, 402-13.
- Moyer CL, Dobbs FC, Karl DM, 1994. Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. *Applied and Environmental Microbiology* **60**, 871-9.
- Mukhopadhyay K, Garrison NK, Hinton DM, *et al.*, 1996. Identification and characterization of bacterial endophytes of rice. *Mycopathologia* **134**, 151-9.
- Müller DB, Vogel C, Bai Y, Vorholt JA, 2016. The Plant Microbiota: Systems-Level Insights and Perspectives. *Annual Review of Genetics* **50**, 211-34.
- Mwajita MR, Murage H, Tani A, Kahangi EM, 2013. Evaluation of rhizosphere, rhizoplane and phyllosphere bacteria and fungi isolated from rice in Kenya for plant growth promoters. *SpringerPlus* **2**, 606-.
- Mysara M, Saeys Y, Leys N, Raes J, Monsieurs P, 2015. CATCh, an Ensemble Classifier for Chimera Detection in 16S rRNA Sequencing Studies. *Applied and Environmental Microbiology* **81**, 1573-84.
- Naik BS, Shashikala J, Krishnamurthy YL, 2009. Study on the diversity of endophytic communities from rice (*Oryza sativa* L.) and their antagonistic activities in vitro. *Microbiol Res* **164**, 290-6.
- Nalley L, Tsiboe F, Durand-Morat A, Shew A, Thoma G, 2016. Economic and environmental impact of rice blast pathogen (*Magnaporthe oryzae*) alleviation in the United States. *PLoS One* **11**.
- Nandi AS, Sengupta B, Sen SP, 2009. Utility of *Rhizobium* in the phyllosphere of crop plants in nitrogen-free sand culture. *The Journal of Agricultural Science* **98**, 167-71.

- Neha KV, Balabaskar P, Ramasamy N, 2016. Survey and occurrence of *Rhizoctonia solani* (Kuhn) causing sheath blight of rice and in vitro efficacy of bacterial antagonists against *Rhizoctonia solani* (Kuhn). **37**, 1421-7.
- Nelson EB, 2018. The seed microbiome: origins, interactions and impacts. *Plant Soil* **422**, 7-34.
- Nelson MC, Morrison HG, Benjamino J, Grim SL, Graf J, 2014. Analysis, Optimization and Verification of Illumina-Generated 16S rRNA Gene Amplicon Surveys. *PLoS One* **9**, e94249.
- Nemergut DR, Schmidt SK, Fukami T, *et al.*, 2013. Patterns and processes of microbial community assembly. *Microbiology and molecular biology reviews : MMBR* **77**, 342-56.
- Niu B, Paulson JN, Zheng X, Kolter R, 2017a. Simplified and representative bacterial community of maize roots. *Proc Natl Acad Sci U S A*. **114**, E2450-E9. doi: 10.1073/pnas.1616148114. Epub 2017 Mar 8.
- Niu B, Paulson JN, Zheng X, Kolter R, 2017b. Simplified and representative bacterial community of maize roots. *Proceedings of the National Academy of Sciences* **114**, E2450-E9.
- Noodén LD, Guamét JJ, John I, 1997. Senescence mechanisms. *Physiologia Plantarum* **101**, 746-53.
- Ofek-Lazar M, Sela N, Goldman-Voronov M, Green SJ, Hadar Y, Minz D, 2014. Niche and host-associated functional signatures of the root surface microbiome. *Nat Commun* **5**, 4950.
- Ohyanagi H, Ebata T, Huang X, *et al.*, 2016. OryzaGenome: Genome Diversity Database of Wild Oryza Species. *Plant & cell physiology* **57**, e1-e.
- Oksanen J, Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Peter, R., Minchin, R., O'hara, R. B., Gavin, L., Solymos, P., Henry, H., Szoecs, E., and Wagner, H., 2018. vegan: Community Ecology Package. *R Package version 2.5.3*.
- Olanrewaju OS, Ayangbenro AS, Glick BR, Babalola OO, 2019. Plant health: feedback effect of root exudates-rhizobiome interactions. *Appl Microbiol Biotechnol* **103**, 1155-66.
- Osmanovic D, Kessler DA, Rabin Y, Soen Y, 2018. Darwinian selection of host and bacteria supports emergence of Lamarckian-like adaptation of the system as a whole. *Biology direct* **13**, 1-13.
- Ottman N, Smidt H, De Vos W, Belzer C, 2012. The function of our microbiota: who is out there and what do they do? *Frontiers in Cellular and Infection Microbiology* **2**.
- Oulhen N, Schulz B, Carrier T, 2016. English translation of Heinrich Anton de Bary's 1878 speech, 'Die Erscheinung der Symbiose' ('De la symbiose'). *Symbiosis* **69**.
- Palanga E, Filloux D, Martin DP, *et al.*, 2016. Metagenomic-Based Screening and Molecular Characterization of Cowpea-Infecting Viruses in Burkina Faso. *PLoS One* **11**, e0165188.
- Pandin C, Le Coq D, Canette A, Aymerich S, Briandet R, 2017. Should the biofilm mode of life be taken into consideration for microbial biocontrol agents? *Microb Biotechnol* **10**, 719-34.
- Park Y-H, Kim Y, Mishra RC, Bae H, 2017. Fungal endophytes inhabiting mountain-cultivated ginseng (*Panax ginseng* Meyer): Diversity and biocontrol activity against ginseng pathogens. *Sci Rep* **7**, 16221.
- Paszkowski U, Kroken S, Roux C, Briggs SP, 2002. Rice phosphate transporters include an evolutionarily divergent gene specifically activated in arbuscular mycorrhizal symbiosis. *Proceedings of the National Academy of Sciences* **99**, 13324-9.
- Peiffer JA, Spor A, Koren O, *et al.*, 2013. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences* **110**, 6548-53.
- Pereira-Marques J, Hout A, Ferreira RM, *et al.*, 2019. Impact of Host DNA and Sequencing Depth on the Taxonomic Resolution of Whole Metagenome Sequencing for Microbiome Analysis. *Front Microbiol* **10**.

- Philippe N, Legendre M, Doutre G, *et al.*, 2013. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* **341**, 281-6.
- Pieterse CM, Zamioudis C, Berendsen RL, Weller DM, Van Wees SC, Bakker PA, 2014. Induced systemic resistance by beneficial microbes. *Annual Review of Phytopathology* **52**, 347-75.
- Pili NN, França SC, Kyndt T, *et al.*, 2016. Analysis of fungal endophytes associated with rice roots from irrigated and upland ecosystems in Kenya. *Plant and Soil* **405**, 371-80.
- Prakobsub K, Ashizawa T, 2017. Intercellular invasion of rice roots at the seedling stage by the rice false smut pathogen, *Villosiclava virens*. *J. Gen. Plant Pathol.* **83**, 358-61.
- Priya H, Prasanna R, Ramakrishnan B, *et al.*, 2015. Influence of cyanobacterial inoculation on the culturable microbiome and growth of rice. *Microbiol Res* **171**, 78-89.
- Qian X, Li H, Wang Y, *et al.*, 2019. Leaf and Root Endospheres Harbor Lower Fungal Diversity and Less Complex Fungal Co-occurrence Patterns Than Rhizosphere. *Front Microbiol* **10**.
- Quince C, Walker A, Simpson J, Loman N, Segata N, 2017. Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology* **35**, 833-44.
- Ramette A, Tiedje JM, 2007. Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. *Proceedings of the National Academy of Sciences* **104**, 2761-6.
- Rampelli S, Soverini M, D'amico F, *et al.*, 2019. Shotgun Metagenomics of Human Gut Microbiota Up to Extreme Longevity and the Increasing Role of Xenobiotics Degradation. *SSRN Electronic Journal*.
- Ranju S, Gowthami Y, Samuel S, Gnanamanickam, Palani P, 2018. Bacteriophages: A New Weapon for the Control of Bacterial Blight Disease in Rice Caused by *Xanthomonas oryzae*. *Microbiology and Biotechnology Letters* **46**, 346-59.
- Raoult D, Audic S, Robert C, *et al.*, 2004. The 1.2-megabase genome sequence of Mimivirus. *Science* **306**, 1344-50.
- Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N, 2010. The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ Microbiol* **12**, 2885-93.
- Regalado J, Lundberg DS, Deusch O, *et al.*, 2019. Combining whole genome shotgun sequencing and rDNA amplicon analyses to improve detection of microbe-microbe interaction networks in plant leaves. *bioRxiv*, 823492.
- Reinhold-Hurek B, Bunker W, Burbano CS, Sabale M, Hurek T, 2015. Roots shaping their microbiome: global hotspots for microbial activity. *Annual Review of Phytopathology* **53:403-24.**, 10.1146/annurev-phyto-082712-102342.
- Reinhold-Hurek B, Hurek T, 2011. Living inside plants: bacterial endophytes. *Current Opinion in Plant Biology* **14**, 435-43.
- Rengel Z, 2002. Genetic control of root exudation. *Plant and Soil* **245**, 59-70.
- Rhoads A, Au KF, 2015. PacBio Sequencing and Its Applications. *Genomics Proteomics Bioinformatics* **13**, 278-89.
- Rinke C, Schwientek P, Sczyrba A, *et al.*, 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature. Nature* **499**.
- Rintala A, Pietilä S, Munukka E, *et al.*, 2017. Gut Microbiota Analysis Results Are Highly Dependent on the 16S rRNA Gene Target Region, Whereas the Impact of DNA Extraction Is Minor. *J Biomol Tech* **28**, 19-30.

- Ritpitakphong U, Falquet L, Vimoltust A, Berger A, Métraux JP, L'haridon F, 2016. The microbiome of the leaf surface of *Arabidopsis* protects against a fungal pathogen. *New Phytologist* **210**, 1033-43.
- Rohwer F, Seguritan V, Azam F, Knowlton N, 2002. Rohwer F, Seguritan V, Azam F, Knowlton N.. Diversity and distribution of coral-associated bacteria. *Mar Ecol Prog Ser* 243: 1-10. *Marine Ecology-progress Series - MAR ECOL-PROGR SER* **243**, 1-10.
- Rolli E, Marasco R, Vigani G, *et al.*, 2015. Improved plant resistance to drought is promoted by the root-associated microbiome as a water stress-dependent trait. *Environ Microbiol* **17**, 316-31.
- Roman-Reyna V, Pinili D, Borjaa FN, *et al.*, 2019. The rice leaf microbiome has a conserved community structure controlled by complex host-microbe interactions. *CELL-HOST-MICROBE-D-19-00340*.
- Roossinck MJ, 2011. The good viruses: viral mutualistic symbioses. *Nature Reviews Microbiology* **9**, 99-108.
- Roossinck MJ, 2015. Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles. *Front Microbiol* **5**, 767-.
- Roossinck MJ, Darren PM, Roumagnac P, 2015. Plant virus metagenomics: advances in virus discovery. *Phytopathology* **105**, 716-27.
- Rosenberg E, Zilber-Rosenberg I, 2011. Symbiosis and development: the hologenome concept. *Birth Defects Res C Embryo Today* **93**, 56-66.
- Rosindell J, Hubbell S, Etienne R, 2011. The Unified Neutral Theory of Biodiversity and Biogeography at Age Ten. *Trends in Ecology and Evolution* **26**, 340-8.
- Roughgarden J, Gilbert SF, Rosenberg E, Zilber-Rosenberg I, Lloyd EA, 2018. Holobionts as units of selection and a model of their population dynamics and evolution. *Biological Theory* **13**, 44-65.
- Rowan-Nash AD, Korry BJ, Mylonakis E, Belenky P, 2019. Cross-Domain and Viral Interactions in the Microbiome. *Microbiology and Molecular Biology Reviews* **83**.
- Sakthivel N, Mortensen C, Mathur S, 2001. Detection of *Xanthomonas oryzae* pv. *oryzae* in artificially inoculated and naturally infected rice seeds and plants by molecular techniques. *Appl Microbiol Biotechnol* **56**, 435-41.
- Samb-Ba B, Mazenot C, Gassama-Sow A, *et al.*, 2014. MALDI-TOF Identification of the Human Gut Microbiome in People with and without Diarrhea in Senegal. *PLoS One* **9**, e87419.
- Santhanam R, Luu VT, Weinhold A, Goldberg J, Oh Y, Baldwin IT, 2015. Native root-associated bacteria rescue a plant from a sudden-wilt disease that emerged during continuous cropping. *Proc Natl Acad Sci U S A* **112**, E5013-E20.
- Santos-Medellín C, Edwards J, Liechty Z, Nguyen B, Sundaresan V, 2017. Drought Stress Results in a Compartment-Specific Restructuring of the Rice Root-Associated Microbiomes. *Mbio* **8**, e00764-17.
- Sasse J, Martinoia E, Northen T, 2018. Feed Your Friends: Do Plant Exudates Shape the Root Microbiome? *Trends Plant Sci* **23**, 25-41.
- Schlaeppli K, Bulgarelli D, 2015. The Plant Microbiome at Work. *Molecular Plant-Microbe Interactions*® **28**, 212-7.
- Schlaeppli K, Dombrowski N, Oter RG, Van Themaat EVL, Schulze-Lefert P, 2014. Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. *Proceedings of the National Academy of Sciences* **111**, 585-92.
- Schloss PD, Handelsman J, 2005. Introducing DOTUR, a Computer Program for Defining Operational Taxonomic Units and Estimating Species Richness. *Applied and Environmental Microbiology* **71**, 1501-6.

- Schloss PD, Westcott SL, Ryabin T, *et al.*, 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology* **75**, 7537-41.
- Schneider GF, Dekker C, 2012. DNA sequencing with nanopores. *Nature Biotechnology* **30**, 326-8.
- Schroeder PJ, Jenkins DG, 2018. How robust are popular beta diversity indices to sampling error? *Ecosphere* **9**, e02100.
- Secor PR, Sweere JM, Michaels LA, *et al.*, 2015. Filamentous Bacteriophage Promote Biofilm Assembly and Function. *Cell Host Microbe* **18**, 549-59.
- Seguritan V, Rohwer F, 2001. FastGroup: A program to dereplicate libraries of 16S rDNA sequences. *BMC Bioinformatics* **2**, 9.
- Sekirov I, Russell SL, Antunes LCM, Finlay BB, 2010. Gut Microbiota in Health and Disease. *Physiol Rev* **90**, 859-904.
- Sessitsch A, Hardoim P, Döring J, *et al.*, 2012. Functional Characteristics of an Endophyte Community Colonizing Rice Roots as Revealed by Metagenomic Analysis. *MPMI* **25**, 28-36.
- Shade A, Jacques M-A, Barret M, 2017. Ecological patterns of seed microbiome diversity, transmission, and assembly. *Current Opinion in Microbiology* **37**, 15-22.
- Shaw A, Mcdaniel S, Werner O, Ros-Espín R, 2002. Invited essay: New frontiers in bryology and lichenology. Phylogeography and phylodemography. *Bryologist* **105**, 373-83.
- Shenton M, Iwamoto C, Kurata N, Ikeo K, 2016. Effect of wild and cultivated rice genotypes on rhizosphere bacterial community composition. *Rice* **9**, 42.
- Shi S, Chang J, Tian L, *et al.*, 2019. Comparative analysis of the rhizomicrobiome of the wild versus cultivated crop: insights from rice and soybean. *Archives of Microbiology* **201**, 879-88.
- Sieber M, Pita L, Weiland-Bräuer N, *et al.*, 2019. Neutrality in the Metaorganism. *PLoS Biology* **17**, e3000298.
- Silva Parra A, Mogollón Ortiz ÁM, Delgado Huertas H, 2018. Soil microbiota: Influence of different land use patterns and soil management factors at Villavicencio Oxisol, East Colombia. *Biota Colombiana* **18**, 1-11.
- Silva UC, Medeiros JD, Leite LR, *et al.*, 2017. Long-Term Rock Phosphate Fertilization Impacts the Microbial Communities of Maize Rhizosphere. *Front Microbiol* **8**.
- Singer E, Bonnette J, Kenaley SC, Woyke T, Juenger TE, 2019. Plant compartment and genetic variation drive microbiome composition in switchgrass roots. *Environmental Microbiology Reports* **11**, 185-95.
- Skamnioti P, Gurr S, 2009. Against the grain: safeguarding rice from rice blast disease. *Trends in biotechnology* **27**, 141-50.
- Somers E, Vanderleyden J, Srinivasan M, 2004. Rhizosphere bacterial signalling: a love parade beneath our feet. *Crit Rev Microbiol* **30**, 205-40.
- Spellerberg I, Fedor P, 2003. A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the ‘Shannon–Wiener’ Index. *Global Ecology & Biogeography* **12**, 177-9.
- Stackebrandt E, Goebel BM, 1994. Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. *International Journal of Systematic and Evolutionary Microbiology* **44**, 846-9.
- Steenhoudt O, Vanderleyden J, 2000. Azospirillum, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. *FEMS Microbiol Rev* **24**, 487-506.

- Stegen JC, Lin X, Konopka AE, Fredrickson JK, 2012. Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME journal* **6**, 1653-64.
- Stringlis IA, Yu K, Feussner K, *et al.*, 2018. MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. *Proceedings of the National Academy of Sciences* **115**, E5213-E22.
- Syranidou E, Christofilopoulos S, Gkavrou G, *et al.*, 2016. Exploitation of Endophytic Bacteria to Enhance the Phytoremediation Potential of the Wetland Helophyte *Juncus acutus*. *Front Microbiol* **7**, 1016.
- Tandina F, Almeras L, Koné AK, Doumbo OK, Raoult D, Parola P, 2016. Use of MALDI-TOF MS and culturomics to identify mosquitoes and their midgut microbiota. *Parasit Vectors* **9**, 495.
- Taur Y, Coyte K, Schluter J, *et al.*, 2018. Reconstitution of the gut microbiota of antibiotic-treated patients by autologous fecal microbiota transplant. *Science Translational Medicine* **10**, eaap9489.
- Thapa S, Prasanna R, Ramakrishnan B, *et al.*, 2018. Interactive effects of Magnaporthe inoculation and nitrogen doses on the plant enzyme machinery and phyllosphere microbiome of resistant and susceptible rice cultivars. *Arch Microbiol*, 1540-0.
- Theis KR, Dheilly NM, Klassen JL, *et al.*, 2016. Getting the hologenome concept right: an eco-evolutionary framework for hosts and their microbiomes. *mSystems* **1**.
- Thijs S, Op De Beeck M, Beckers B, *et al.*, 2017. Comparative Evaluation of Four Bacteria-Specific Primer Pairs for 16S rRNA Gene Surveys. *Front Microbiol* **8**.
- Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F, 2009. Laboratory procedures to generate viral metagenomes. *Nat Protoc* **4**, 470-83.
- Tian L, Wu A-K, Friedman J, Waldor MK, Weiss ST, Liu Y-Y, 2017. Deciphering Functional Redundancy in the Human Microbiome. *bioRxiv*, 176313.
- Tkacz A, Poole P, 2015. Role of root microbiota in plant productivity. *J Exp Bot* **66**, 2167-75.
- Truyens S, Weyens N, Cuypers A, Vangronsveld J, 2015. Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environmental Microbiology Reports* **7**, 40-50.
- Tshikantwa TS, Ullah MW, He F, Yang G, 2018. Current Trends and Potential Applications of Microbial Interactions for Human Welfare. *Front Microbiol* **9**, 1156-.
- Turner TR, James EK, Poole PS, 2013. The plant microbiome. *Genome Biology* **14**, 209.
- Uehara-Ichiki T, Shiba T, Matsukura K, Ueno T, Hirae M, Sasaya T, 2013. Detection and diagnosis of rice-infecting viruses. *Front Microbiol* **4**, 289-.
- Urayama S, Kato S, Suzuki Y, *et al.*, 2010. Mycoviruses related to chrysovirus affect vegetative growth in the rice blast fungus *Magnaporthe oryzae*. *J Gen Virol* **91**, 3085-94.
- Urayama S, Ohta T, Onozuka N, *et al.*, 2012. Characterization of *Magnaporthe oryzae* chrysovirus 1 structural proteins and their expression in *Saccharomyces cerevisiae*. *Journal of Virology* **86**, 8287-95.
- Vacher C, Hampe A, Porté A, Sauer U, Compant S, Morris C, 2016a. The Phyllosphere: Microbial Jungle at the Plant-Climate Interface. *Annual Review of Ecology Evolution and Systematics* **47**.
- Vacher C, Tamaddoni-Nezhad A, Kamenova S, *et al.*, 2016b. Chapter One - Learning Ecological Networks from Next-Generation Sequencing Data. In: Woodward G, Bohan DA, eds. *Advances in Ecological Research*. Academic Press, 1-39. (54.)
- Valk T, Vezzi F, Ormestad M, Dalén L, Guschanski K, 2019. Index hopping on the Illumina HiSeqX platform and its consequences for ancient DNA studies. *Molecular Ecology Resources*.

- Valverde RA, Sabanadzovic S, Rush MC, 2011. Identification of *Oryza sativa* endornavirus in rice genotypes from breeding programmes in the United States. *Plant Breeding* **130**, 271-4.
- Van Der Ent S, Van Hulten M, Pozo MJ, *et al.*, 2009. Priming of plant innate immunity by rhizobacteria and β -aminobutyric acid: differences and similarities in regulation. *New Phytologist* **183**, 419-31.
- Van Der Ent S, Verhagen BW, Van Doorn R, *et al.*, 2008. MYB72 is required in early signaling steps of rhizobacteria-induced systemic resistance in *Arabidopsis*. *Plant Physiology* **146**, 1293-304.
- Van Der Heijden M, Bruin S, Luckerhoff L, Logtestijn R, Schlaeppi K, 2015. A widespread plant-fungal-bacterial symbiosis promotes plant biodiversity, plant nutrition and seedling recruitment. *The ISME journal* **10**.
- Van Elsas JD, Duarte GF, Rosado AS, Smalla K, 1998. Microbiological and molecular biological methods for monitoring microbial inoculants and their effects in the soil environment. *J Microbiol Methods* **32**, 133-54.
- Vandenkoornhuysen P, Quaiser A, Duhamel M, Le Van A, Dufresne A, 2015. The importance of the microbiome of the plant holobiont. *New Phytol* **206**, 1196-206.
- Vannier N, Agler M, Hacquard S, 2019. Microbiota-mediated disease resistance in plants. *PLoS Pathogens* **15**.
- Vannier N, Mony C, Bittebiere A-K, Michon-Coudouel S, Biget M, Vandenkoornhuysen P, 2018. A microorganisms' journey between plant generations. *Microbiome* **6**, 79.
- Vellend M, 2010. Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology* **85**, 183-206.
- Verma SK, Kingsley K, Irizarry I, Bergen M, Kharwar RN, White JF, Jr., 2017. Seed-vectored endophytic bacteria modulate development of rice seedlings. *Journal of Applied Microbiology* **122**, 1680-91.
- Vermeiren H, Willems A, Schoofs G, *et al.*, 1999. The rice inoculant strain *Alcaligenes faecalis* A15 is a nitrogen-fixing *Pseudomonas stutzeri*. *Syst Appl Microbiol* **22**, 215-24.
- Vetrovsky T, Baldrian P, 2013. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One* **8**, e57923.
- Villanueva-Millán MJ, Pérez-Matute P, Oteo JA, 2015. Gut microbiota: a key player in health and disease. A review focused on obesity. *Journal of Physiology and Biochemistry* **71**, 509-25.
- Villarreal L, 2009. *Origin of Group Identity: Viruses, Addiction and Cooperation*.
- Vorholt JA, 2012. Microbial life in the phyllosphere. *Nature Rev. Microbiol.* **10**, 828-40.
- Wagner BD, Grunwald GK, Zerbe GO, *et al.*, 2018. On the Use of Diversity Measures in Longitudinal Sequencing Studies of Microbial Communities. *Front Microbiol* **9**, 1037-.
- Wagner J, Coupland P, Browne HP, Lawley TD, Francis SC, Parkhill J, 2016. Evaluation of PacBio sequencing for full-length bacterial 16S rRNA gene classification. *BMC Microbiol* **16**, 274.
- Walitang DI, Kim C-G, Kim K, Kang Y, Kim YK, Sa T, 2018. The influence of host genotype and salt stress on the seed endophytic community of salt-sensitive and salt-tolerant rice cultivars. *BMC Plant Biology* **18**, 51.
- Wallace J, Laforest-Lapointe I, Kembel S, 2018. Variation in the leaf and root microbiome of sugar maple (*Acer saccharum*) at an elevational range limit. *PeerJ* **6**, e5293.
- Wang DG, Fan J-B, Siao C-J, *et al.*, 1998. Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome. *Science* **280**, 1077-82.
- Wang GCY, Wang Y, 1996. The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. *Microbiology* **142**, 1107-14.

- Wang W, Mauleon R, Hu Z, *et al.*, 2018. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* **557**, 43-9.
- Wang W, Zhai Y, Cao L, Tan H, Zhang R, 2016. Endophytic bacterial and fungal microbiota in sprouts, roots and stems of rice (*Oryza sativa* L.). *Microbiol Res* **188-189**, 1-8.
- Weiss S, Xu ZZ, Peddada S, *et al.*, 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5**, 27-.
- Westcott SL, Schloss PD, 2017. OptiClust, an Improved Method for Assigning Amplicon-Based Sequence Data to Operational Taxonomic Units. *mSphere* **2**, e00073-17.
- Whittaker RH, 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* **30**, 279-338.
- Wiebe K. How to feed the world in 2050. *Proceedings of the Insights from an expert meeting at FAO, 2009*, 24-6.
- Willis C, Desai D, Laroche J, 2019. Influence of 16S rRNA variable region on perceived diversity of marine microbial communities of the Northern North Atlantic. *FEMS Microbiology Letters* **366**, fnz152.
- Wischmeyer PE, Mcdonald D, Knight R, 2016. Role of the microbiome, probiotics, and 'dysbiosis therapy' in critical illness. *Curr Opin Crit Care* **22**, 347-53.
- Woese CR, Fox GE, 1977. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences* **74**, 5088-90.
- Wu L, Liu R, Niu Y, *et al.*, 2016. Whole genome sequence of *Pantoea ananatis* R100, an antagonistic bacterium isolated from rice seed. *Journal of Biotechnology* **225**, 1-2.
- Wu S, Manber U. Agrep—a fast approximate pattern-matching tool. *Proceedings of the Usenix Winter 1992 Technical Conference, 1992*, 153-62.
- Xia Y, Debolt S, Dreyer J, Scott D, Williams MA, 2015. Characterization of culturable bacterial endophytes and their capacity to promote plant growth from plants grown using organic or conventional practices. *Front Plant Sci* **6**.
- Xia Y, Sun J, Chen D-G, 2018. *Statistical analysis of microbiome data with R*. Springer.
- Xu T, Li Y, Zeng X, *et al.*, 2017. Isolation and evaluation of endophytic *Streptomyces endus* OsiSh-2 with potential application for biocontrol of rice blast disease. *J Sci Food Agric* **97**, 1149-57.
- Xu X-H, Su Z-Z, Wang C, *et al.*, 2014. The rice endophyte *Harpophora oryzae* genome reveals evolution from a pathogen to a mutualistic endophyte. *Sci Rep* **4**, 5783.
- Yadav M, S A, Lal DB, Gautam P, Ghritlahre S, 2015. Mycoviruses: An Introduction to Potential Biocontrol Agents. *Popular kheti*, 68070.
- Yan H, Yu S, Xie G-L, Fang W, Su T, Li B, 2010. Grain Discoloration of Rice Caused by *Pantoea ananatis* (synonym *Erwinia uredovora*) in China. *Plant Disease - PLANT DIS* **94**, 482-.
- Yang B, Wang Y, Qian P-Y, 2016. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* **17**.
- Yang R-H, Su J-H, Shang J-J, *et al.*, 2018. Evaluation of the ribosomal DNA internal transcribed spacer (ITS), specifically ITS1 and ITS2, for the analysis of fungal diversity by deep sequencing. *PLoS One* **13**, e0206428.
- Yarza P, Yilmaz P, Pruesse E, *et al.*, 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* **12**, 635-45.
- Yong M, Liu Y, Chen T, Fan L, Wang Z, Hu D, 2018. Cytological studies on the infection of rice root by *Ustilagoidea virens*. *Microsc Res Tech* **81**, 389-96.

- Yuan ZL, Su ZZ, Mao LJ, *et al.*, 2011. Distinctive endophytic fungal assemblage in stems of wild rice (*Oryza granulata*) in China with special reference to two species of *Muscador* (Xylariaceae). *J Microbiol* **49**, 15-23.
- Zahn G, Amend AS, 2017. Foliar microbiome transplants confer disease resistance in a critically-endangered plant. *PeerJ* **5**, e4020.
- Zamioudis C, Korteland J, Van Pelt JA, *et al.*, 2015. Rhizobacterial volatiles and photosynthesis-related signals coordinate MYB 72 expression in *Arabidopsis* roots during onset of induced systemic resistance and iron-deficiency responses. *The Plant Journal* **84**, 309-22.
- Zhang A, Lu Q, Yin Y, Ding S, Wen X, Lu C, 2010a. Comparative proteomic analysis provides new insights into the regulation of carbon metabolism during leaf senescence of rice grown under field conditions. *J Plant Physiol* **167**, 1380-9.
- Zhang H-M, Yang J, Chen J-P, Adams MJ, 2008. A black-streaked dwarf disease on rice in China is caused by a novel fijivirus. *Archives of Virology* **153**, 1893-8.
- Zhang J, Zhang C, Yang J, *et al.*, 2019. Insights into Endophytic Bacterial Community Structures of Seeds Among Various *Oryza sativa* L. Rice Genotypes. *Journal of Plant Growth Regulation* **38**, 93-102.
- Zhang JY, Zhang N, Liu YX, *et al.*, 2018. Root microbiota shift in rice correlates with resident time in the field and developmental stage. *Science China-Life Sciences* **61**, 613-21.
- Zhang L, Birch RG, 1997. The gene for albicidin detoxification from *Pantoea dispersa* encodes an esterase and attenuates pathogenicity of *Xanthomonas albilineans* to sugarcane. *Proc Natl Acad Sci U S A* **94**, 9984-9.
- Zhang L, Gao JS, Kim SG, *et al.*, 2016. *Novosphingobium oryzae* sp. nov., a potential plant-promoting endophytic bacterium isolated from rice roots. *International Journal of Systematic and Evolutionary Microbiology* **66**, 302-7.
- Zhang X, Li E, Xiong X, Shen D, Feng Y, 2010b. Colonization of endophyte *Pantoea agglomerans* YS19 on host rice, with formation of multicellular symplasmata. *World Journal of Microbiology and Biotechnology* **26**, 1667-73.
- Zhao J, Gao Q, Zhou J, *et al.*, 2019. The scale dependence of fungal community distribution in paddy soil driven by stochastic and deterministic processes. *Fungal Ecology* **42**, 100856.
- Zheng L, Shu C, Zhang M, Yang M, Zhou E, 2019. Molecular Characterization of a Novel Endornavirus Conferring Hypovirulence in Rice Sheath Blight Fungus *Rhizoctonia solani* AG-1 IA Strain GD-2. *Viruses* **11**.
- Zhou J, Deng Y, Zhang P, *et al.*, 2014. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proceedings of the National Academy of Sciences* **111**, E836-E45.
- Zhou J, Kang S, Schadt CW, Garten CT, 2008. Spatial scaling of functional gene diversity across various microbial taxa. *Proceedings of the National Academy of Sciences* **105**, 7768-73.
- Zhou J, Ning D, 2017. Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiology and Molecular Biology Reviews* **81**.
- Zhu Y, Chen H, Fan J, *et al.*, 2000. Genetic diversity and disease control in rice. *Nature*. **406**, 718-22.
- Zou Q, Lin G, Jiang X, Liu X, Zeng X, 2018. Sequence clustering in bioinformatics: an empirical study. *Briefings in Bioinformatics* **21**, 1-10.

Annexes

Annexe 1 :

**Tableau récapitulatif des OTUs différentiellement abondantes
entre les compartiments de la plante à l'issue du test des rangs
signés de Wilcoxon sur les abondances relatives des OTUs dans
chaque compartiment végétal**

Graine abondance relative	Racine abondance relative	Tige abondance relative	Phylum	Class	Order	Family	Genus	median_diff	mean_diff	wilcox_p_value
6875.2	1160.1	7945.1	Proteobacteria	Alphaproteobacteria	Rickettsiales	Mitochondria	NA	7388.5	5715.2	4.04E-37
1288.6	144.2	2060.5	Cyanobacteria	Oxyphotobacteria	Chloroplast	NA	NA	1361.5	1144.4	2.22E-29
0.0	1200.9	31.4	Epsilonbacteraeota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	-446.0	-1200.9	2.31E-36
687.4	0.5	31.2	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Pantoea	7.5	686.9	3.75E-35
635.3	0.3	29.5	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	NA	7.5	635.1	1.55E-35
0.0	283.6	3.0	Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Azospirillum	-65.0	-283.6	3.26E-31
352.4	0.3	0.6	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	3.0	352.1	1.32E-27
0.0	188.0	7.3	Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Niveispirillum	-36.0	-188.0	9.09E-30
0.0	143.8	0.5	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	-21.0	-143.8	1.57E-29
0.0	141.9	0.3	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Azospira	-31.0	-141.9	1.04E-32
0.0	105.8	0.1	Epsilonbacteraeota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	-4.0	-105.8	1.73E-17
0.0	75.2	0.1	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Sporomusa	-2.0	-75.2	7.96E-17
0.0	93.0	7.1	Bacteroidetes	Bacteroidia	Bacteroidales	Paludibacteraceae	Paludibacter	-8.0	-93.0	3.11E-21
0.0	68.9	0.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Propionivibrio	-5.0	-68.9	9.50E-20
0.0	59.1	1.0	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	-7.5	-59.1	1.76E-20
0.0	47.5	7.3	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Spirochaeta_2	-10.0	-47.5	2.65E-29
0.0	64.0	0.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	NA	-6.0	-64.0	7.44E-19
0.0	43.9	0.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Acidovorax	-6.0	-43.9	3.27E-24
0.0	54.8	0.8	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-14.0	-54.8	1.12E-27
0.0	44.4	0.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Hydrogenophaga	-9.0	-44.4	4.01E-26
0.0	50.0	0.4	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-3.0	-50.0	1.65E-18
0.0	51.2	0.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	NA	-5.0	-51.2	2.14E-22
0.0	49.4	1.6	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Uliginosibacterium	-2.0	-49.4	7.96E-17
0.0	52.4	1.0	Proteobacteria	Alphaproteobacteria	Rhizobiales	Pleomorphomonadaceae	Ancalomicrobium	-4.5	-52.4	4.93E-19
0.0	46.6	1.1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Magnetospirillaceae	Magnetospirillum	-4.0	-46.6	2.66E-20
0.0	29.7	0.1	Firmicutes	Clostridia	Clostridiales	Family_XII	Acidaminobacter	-4.5	-29.7	6.21E-20
0.0	40.1	0.2	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-3.0	-40.1	1.73E-17
0.0	32.1	0.6	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Treponema	-2.0	-32.1	7.96E-17
0.0	36.9	0.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Candidatus_Accumulibacter	-3.0	-36.9	2.08E-23
0.0	33.0	0.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	NA	-9.0	-33.0	8.68E-27
0.0	34.4	0.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Azovibrio	-2.0	-34.4	1.69E-16
0.0	31.1	0.2	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Acidovorax	-3.5	-31.1	2.06E-23
0.0	26.1	0.1	Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Niveispirillum	-2.0	-26.1	6.15E-20
0.0	28.2	0.1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Novispirillum	-1.0	-28.2	4.47E-15

0.0	37.8	0.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Propionivibrio	-5.0	-37.8	5.10E-27
0.0	19.6	2.0	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	-3.5	-19.6	1.13E-20
0.0	17.2	0.0	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-1.0	-17.2	2.43E-16
0.0	20.2	0.1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-1.0	-20.2	2.35E-14
0.0	17.3	0.6	Epsilonbacteraeota	Campylobacteria	Campylobacteriales	Thiovulaceae	Sulfuricurvum	-6.0	-17.3	3.28E-24
0.0	24.2	0.4	Proteobacteria	Alphaproteobacteria	Rhizobiales	Pleomorphomonadaceae	NA	-1.0	-24.2	1.82E-14
0.0	7.9	0.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Hydrogenophaga	-1.0	-7.9	1.59E-17
0.0	20.1	0.1	Proteobacteria	Deltaproteobacteria	Myxococcales	Archangiaceae	Anaeromyxobacter	-1.0	-20.1	7.93E-17
0.0	12.1	0.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Hydrogenophaga	-2.0	-12.1	2.60E-20
0.0	12.5	6.4	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	-1.0	-12.5	1.55E-15
0.0	23.3	0.3	Proteobacteria	Alphaproteobacteria	Elsterales	Elsteraceae	Elstera	-1.0	-23.3	4.47E-15
0.0	17.3	0.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Ideonella	-3.0	-17.3	7.28E-21
0.0	19.6	0.0	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-4.0	-19.6	1.18E-28
0.0	14.5	1.9	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-1.0	-14.5	1.15E-16
0.0	15.5	0.1	Verrucomicrobia	Verrucomicrobiae	Opitutales	Opitutaceae	Lacunisphaera	-6.0	-15.5	5.18E-27
0.0	13.4	0.2	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Sulfuritalea	-2.5	-13.4	4.69E-21
0.0	15.0	1.0	Proteobacteria	Alphaproteobacteria	Reyranellales	Reyranellaceae	Reyranella	-2.0	-15.0	3.22E-19
0.0	12.2	0.4	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Salinispira	-2.0	-12.2	7.24E-19
0.0	10.7	0.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Curvibacter	-2.0	-10.7	1.66E-20
0.0	11.2	0.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Propionivibrio	-3.0	-11.2	3.04E-21
0.0	3.0	0.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Hydrogenophaga	-1.0	-3.0	7.30E-17
0.0	14.1	0.0	Acidobacteria	Holophagae	Holophagales	Holophagaceae	Geothrix	-1.0	-14.1	1.07E-15
0.0	12.3	0.3	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	NA	-1.0	-12.3	1.54E-15
0.0	12.6	0.1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-1.0	-12.6	1.15E-16
0.0	11.2	0.8	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	NA	-2.0	-11.2	1.08E-18
0.0	9.2	0.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	NA	-2.0	-9.2	6.00E-20
0.0	11.8	0.0	Spirochaetes	Leptospirae	Leptospirales	Leptospiraceae	Leptonema	-1.0	-11.8	1.07E-15
0.0	9.9	0.2	Proteobacteria	Alphaproteobacteria	Dongiales	Dongiaceae	Dongia	-1.0	-9.9	3.51E-16
0.0	7.6	0.7	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Candidatus_Accumulibacter	-1.0	-7.6	7.81E-17
0.0	9.4	0.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	NA	NA	-1.0	-9.4	2.52E-14
0.0	6.9	0.9	Actinobacteria	Actinobacteria	Micrococcales	Cellulomonadaceae	Pseudactinotalea	-0.5	-6.9	3.57E-14
0.0	7.8	0.6	Verrucomicrobia	Verrucomicrobiae	Opitutales	Opitutaceae	Lacunisphaera	-3.0	-7.8	1.25E-21
0.0	7.0	0.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	NA	-0.5	-7.0	3.49E-14
0.0	8.4	0.1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-1.0	-8.4	2.40E-16
0.0	6.5	0.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Dechloromonas	-1.0	-6.5	6.14E-15
0.0	6.6	0.2	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	NA	-0.5	-6.6	3.57E-14
0.0	1.8	0.0	Epsilonbacteraeota	Campylobacteria	Campylobacteriales	Sulfurospirillaceae	Sulfurospirillum	-1.0	-1.8	4.61E-16

Annexe 2 :

**Tableau récapitulatif des OTUs racinaires différentiellement abondantes
entre les stades de développement du riz à l'issue du test des rangs signés de
Wilcoxon sur les abondances relatives des OTUs chaque date de
prélèvements**

Abundance relative moyenne floraison 2017	Abundance relative moyenne maturation 2017	Abundance relative moyenne tallage 2018	Abundance relative moyenne floraison 2018	Abundance relative moyenne maturation 2018	Phylum	Class	Order	Family	Genus	median_diff	mean_diff	wilcox_p_value
932.8	352.1	771.8	281.8	495.1	Epsilonbacteraeota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	-510.0	-580.7	3.8E-03
163.3	154.7	110.5	103.6	169.7	Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Azospirillum	-46.0	-59.8	5.3E-03
121.2	165.0	65.2	51.7	87.9	Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Niveispirillum	29.0	113.3	3.7E-02
0.0	13.9	0.1	216.3	8.7	Proteobacteria	Deltaproteobacteria	Myxococcales	NA	NA	-8.0	-202.4	1.3E-03
158.5	52.3	91.4	52.9	61.6	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	-37.5	-106.2	4.8E-03
49.6	79.8	69.7	68.1	57.2	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Azospira	21.0	11.7	1.6E-02
99.1	5.3	76.1	3.4	70.7	Epsilonbacteraeota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	-12.0	-93.8	1.2E-05
170.5	6.8	37.9	0.0	0.5	Proteobacteria	Gammaproteobacteria	NA	NA	NA	-22.5	-163.7	4.7E-08
6.8	14.6	86.5	6.3	136.5	Chloroflexi	Chloroflexia	Chloroflexales	Roseiflexaceae	NA	-3.5	-71.9	2.0E-02
76.3	9.9	17.8	18.9	53.9	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Sporomusa	-6.0	-66.4	3.9E-04
0.7	61.3	134.4	34.4	59.4	Bacteroidetes	Bacteroidia	Bacteroidales	Paludibacteraceae	Paludibacter	10.5	60.6	6.5E-07
15.9	10.0	28.3	80.2	19.5	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Propionivibrio	-4.0	-5.9	1.7E-02
0.5	60.7	64.5	100.9	11.6	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	NA	4.0	60.2	2.3E-06
50.0	4.1	20.8	1.1	2.1	Proteobacteria	Gammaproteobacteria	NA	NA	NA	-6.0	-46.0	2.7E-05
36.4	20.1	58.0	8.8	21.0	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	-14.5	-37.8	3.6E-04
0.1	5.4	9.7	55.8	52.9	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	NA	-1.0	-52.8	8.6E-06
5.7	57.1	15.6	32.7	26.6	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Spirochaeta_2	9.0	51.4	3.6E-03
12.1	47.2	18.4	29.8	44.2	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	NA	-11.0	-14.4	9.4E-03
18.9	19.4	61.7	2.7	16.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Acidovorax	-16.0	-42.3	1.2E-04
24.4	35.2	73.9	1.3	9.3	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	-1.0	10.7	1.6E-03
28.4	56.3	20.9	27.6	10.5	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	11.5	45.8	2.1E-02
36.3	40.5	13.6	8.2	9.9	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Hydrogenophaga	16.0	32.2	9.3E-05
1.4	73.0	13.4	40.8	4.0	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	8.0	71.7	3.2E-04
32.5	29.8	22.0	12.4	41.5	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	NA	-8.5	7.8	3.6E-03
7.7	29.8	46.7	20.3	23.3	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Uliginosibacterium	-8.5	-26.4	1.2E-03
57.0	9.8	27.0	2.6	6.2	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Sporomusa	-2.0	-47.3	2.0E-03
12.8	12.4	22.5	14.1	62.7	Proteobacteria	Alphaproteobacteria	Rhizobiales	Pleomorphomonadaceae	Ancalomicrobium	-6.0	-10.2	4.2E-03
10.2	29.2	41.3	9.4	16.6	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Magnetospirillaceae	Magnetospirillum	-6.5	-31.9	2.8E-03
9.6	3.2	74.8	3.0	11.6	Bacteroidetes	Bacteroidia	Bacteroidales	Paludibacteraceae	Paludibacter	-3.5	-71.6	1.7E-05
3.7	10.4	25.7	21.4	36.5	Proteobacteria	Deltaproteobacteria	Desulfuromonadales	Desulfuromonadaceae	Pelobacter	-4.5	-15.2	2.8E-06
29.5	16.8	23.4	1.9	2.5	Firmicutes	Clostridia	Clostridiales	Family_XII	Acidaminobacter	-13.0	-12.7	1.0E-04
52.2	23.8	12.9	8.9	5.3	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-10.0	-43.3	2.3E-04
0.3	5.8	3.8	30.5	38.5	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Treponema	-4.0	-24.7	2.0E-02
0.0	5.5	0.0	65.8	17.5	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	NA	-4.0	-60.4	7.2E-06
3.8	8.6	18.8	18.3	46.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Candidatus_Accumulibacter	-3.0	-10.2	3.9E-05
0.1	1.4	18.0	5.1	48.4	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	NA	-2.0	-16.6	1.3E-05
0.4	4.7	40.8	3.3	23.6	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	NA	-1.0	-36.1	2.0E-03
0.1	10.7	0.4	75.9	11.7	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	Methylocystis	-9.0	-65.2	2.4E-04
2.3	1.9	17.2	7.5	25.7	Bacteroidetes	Bacteroidia	Bacteroidales	Paludibacteraceae	NA	-1.0	-23.7	9.9E-06
4.3	56.4	8.2	21.7	9.1	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	NA	21.0	52.1	1.2E-05
12.9	8.1	16.2	5.1	39.5	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Azovibrio	1.0	-4.7	9.3E-03
4.4	58.0	6.2	12.0	11.3	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Acidovorax	9.0	46.0	4.0E-02
32.4	9.3	25.7	9.8	17.7	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Novispirillum	-2.5	-22.5	4.0E-07
2.1	5.9	52.4	9.5	18.6	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Propionivibrio	-5.5	-12.7	1.9E-02

Annexe 2

0.9	5.6	3.1	6.3	53.3	Bacteroidetes	Bacteroidia	Bacteroidales	Paludibacteraceae	NA	-9.0	-47.7	4.1E-07
1.5	15.7	10.7	7.6	27.6	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	3.0	14.3	8.3E-03
46.9	0.3	0.7	0.3	0.1	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	-0.5	-46.6	4.7E-06
26.6	4.8	1.9	8.2	1.7	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-3.0	-21.8	1.9E-02
3.1	21.4	11.3	5.8	8.0	Epsilonbacteraeota	Campylobacteria	Campylobacteriales	Thiovulaceae	Sulfuricurvum	11.0	15.7	3.3E-04
1.3	1.6	16.6	13.5	21.7	Proteobacteria	Alphaproteobacteria	Rhizobiales	Pleomorphomonadaceae	NA	-4.0	-15.1	3.1E-07
0.0	13.5	0.6	33.4	5.9	Proteobacteria	Deltaproteobacteria	Myxococcales	Archangiaceae	Anaeromyxobacter	4.5	13.5	1.3E-08
2.8	6.8	4.2	6.5	16.5	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	-1.0	-1.4	1.5E-03
0.7	38.4	10.1	15.9	1.3	Proteobacteria	Alphaproteobacteria	Elsterales	Elsteraceae	Elstera	7.0	22.4	3.7E-03
5.9	13.2	6.5	9.4	12.6	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Ideonella	4.5	7.3	5.7E-04
0.0	10.0	11.0	3.2	12.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Uliginosibacterium	1.0	6.7	1.3E-02
1.9	3.2	23.5	4.4	14.6	Proteobacteria	Deltaproteobacteria	Desulfuromonadales	Geobacteraceae	Geobacter	-5.0	-20.3	8.7E-05
1.8	27.6	1.5	14.6	0.5	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Methylophilaceae	Methylophilus	0.5	25.8	9.4E-03
0.0	13.8	4.6	26.3	7.5	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	NA	2.0	26.3	5.4E-07
3.7	12.3	2.9	19.5	13.3	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	3.0	8.7	4.8E-02
30.3	5.4	4.0	0.1	0.2	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Novispirillum	-2.0	-24.9	2.0E-06
1.4	11.0	7.0	11.0	10.3	Verrucomicrobia	Verrucomicrobiae	Opitutales	Opitutaceae	Lacunisphaera	2.5	9.6	9.3E-03
0.1	19.7	6.2	12.7	8.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Sulfuritalea	4.5	19.7	1.7E-08
1.1	7.2	7.0	10.7	11.4	Proteobacteria	Alphaproteobacteria	Reyranelles	Reyranelleaceae	Reyranelle	-2.5	-4.2	1.7E-02
0.2	0.3	2.7	4.4	19.3	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Treponema	-4.0	-19.0	1.4E-10
10.6	3.4	13.6	1.6	9.4	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Magnetospirillaceae	Magnetospirillum	-1.5	-7.3	3.5E-02
0.2	5.1	3.8	16.5	17.2	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Salinispira	-1.0	-3.6	8.9E-06
30.8	4.0	5.1	2.3	0.6	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Dechloromonas	-7.0	-26.8	6.3E-07
3.4	16.4	6.4	2.3	2.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Curvibacter	7.0	14.1	2.0E-05
7.3	6.5	13.3	0.0	0.7	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Chromobacteriaceae	Vogesella	-1.0	-0.8	1.1E-03
7.2	6.9	5.4	2.8	6.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Propionivibrio	0.5	4.1	2.5E-02
9.0	4.9	17.5	3.6	6.4	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Sporomusa	-3.5	-12.6	2.0E-06
0.3	16.7	3.6	3.5	8.5	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	-1.0	-8.2	2.4E-04
0.1	5.9	1.6	19.6	10.0	Acidobacteria	Holophagae	Holophagales	Holophagaceae	Geothrix	2.0	19.5	3.6E-06
0.4	9.8	2.6	4.1	12.9	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	-1.0	-12.5	2.4E-05
0.5	28.2	3.9	5.5	5.4	Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Niveispirillum	2.5	22.7	2.0E-02
0.3	1.1	2.9	20.8	7.8	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Treponema	-1.5	-6.7	2.2E-08
19.0	0.2	2.1	0.1	5.3	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-3.0	-18.9	4.6E-11
0.0	10.9	0.8	10.4	3.6	Proteobacteria	Gammaproteobacteria	Ectothiorhodospirales	Thioalkalispiraceae	Thioalkalispira	1.0	10.9	1.2E-06
2.3	5.3	6.6	6.6	9.1	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	NA	-2.0	-1.3	3.2E-03
0.0	7.9	1.9	16.2	5.3	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	NA	-2.0	-8.4	1.4E-03
0.4	18.3	2.4	6.3	2.7	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	NA	5.5	12.0	2.4E-03
0.0	10.7	3.2	6.2	11.3	Spirochaetes	Leptospirae	Leptospirales	Leptospiraceae	Leptonema	1.0	10.7	1.5E-06
0.3	8.0	3.7	9.9	22.9	Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Azospirillum	-1.0	-3.4	8.6E-05
0.0	13.5	0.9	7.5	2.1	Bacteroidetes	Bacteroidia	Bacteroidales	Prolixibacteraceae	Roseimarinus	4.0	13.5	3.9E-08
0.0	4.0	1.7	11.9	5.3	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	NA	-1.0	-5.3	2.2E-07
0.7	2.9	3.5	11.8	7.0	Proteobacteria	Alphaproteobacteria	Dongiales	Dongiaceae	Dongia	-1.0	-4.0	2.2E-02
0.0	2.3	1.4	11.0	13.3	Actinobacteria	Actinobacteria	Corynebacteriales	Mycobacteriaceae	Mycobacterium	-2.0	-11.0	7.2E-07
0.0	9.1	0.0	21.2	2.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Gallionellaceae	Sideroxydans	4.0	21.2	1.6E-09
1.6	4.6	4.1	4.9	5.7	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Candidatus_Accumulibacter	-1.0	-2.4	3.7E-03
3.7	4.7	9.2	0.9	5.6	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Sporomusa	-2.0	-4.5	1.4E-06
0.1	19.8	1.6	5.6	1.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	NA	NA	2.5	19.7	4.0E-07
0.2	3.1	2.6	5.6	5.6	Actinobacteria	Actinobacteria	Micrococcales	Cellulomonadaceae	Pseudactinotalea	-2.0	-2.6	3.8E-03
1.0	6.5	2.7	6.0	4.5	Verrucomicrobia	Verrucomicrobiae	Opitutales	Opitutaceae	Lacunisphaera	2.5	5.6	6.9E-04
0.2	7.3	0.2	7.6	6.2	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	NA	2.5	7.1	6.8E-06

Annexe 2

6.0	1.7	5.6	0.0	8.4	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Chromobacteriaceae	Vogesella	-1.0	-4.2	2.5E-06
5.1	3.4	1.5	0.4	1.0	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	-1.0	-1.7	1.5E-02
0.2	2.7	2.4	6.9	5.3	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	-2.0	-2.6	1.1E-02
2.1	6.0	2.6	2.7	1.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Dechloromonas	-1.0	0.1	5.8E-06
7.5	3.1	6.7	0.6	1.7	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	-1.0	-4.4	1.3E-02
1.0	12.9	0.3	2.7	0.3	Bacteroidetes	Bacteroidia	Bacteroidales	Paludibacteraceae	Paludibacter	2.5	10.2	9.7E-03
0.3	3.2	3.8	3.4	2.6	Bacteroidetes	Bacteroidia	Bacteroidales	Prolixibacteraceae	NA	-1.0	-3.6	1.1E-05
0.1	0.5	1.8	5.0	6.5	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Spirochaeta_2	-0.5	-6.0	1.1E-05
0.5	1.0	0.5	2.3	2.2	Verrucomicrobia	Verrucomicrobiae	Opitutales	Opitutaceae	Opitutus	-0.5	-1.3	6.7E-03
0.0	7.9	0.1	5.2	2.7	Proteobacteria	Deltaproteobacteria	Myxococcales	Archangiaceae	Anaeromyxobacter	-0.5	5.2	1.4E-02
5.8	2.2	0.3	0.0	0.0	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Aquabacterium	-0.5	-3.6	1.2E-03
5.5	0.8	4.6	0.1	1.4	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Sporomusa	-0.5	-3.8	1.8E-04
1.6	0.9	10.3	0.0	0.4	Firmicutes	Clostridia	Clostridiales	Family_XII	Acidaminobacter	-0.5	-9.5	3.9E-06
0.0	0.2	0.1	7.4	1.9	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	NA	-1.0	-1.6	1.8E-05
0.7	0.9	2.5	1.7	3.4	Firmicutes	Clostridia	Clostridiales	Family_XIII	Anaerovorax	-1.0	-1.6	3.5E-04
2.8	4.4	4.0	5.6	3.9	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	NA	-0.5	1.5	3.9E-02
0.3	0.2	1.0	0.8	4.9	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Spirochaeta_2	-1.0	-4.7	3.2E-07
4.3	0.6	2.1	0.2	0.6	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Dechloromonas	-2.0	-3.7	2.3E-06
2.0	0.4	2.0	0.4	1.2	Epsilonbacteraeota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	-1.0	-1.5	8.4E-05
1.6	0.6	1.2	0.3	0.7	Epsilonbacteraeota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	-1.0	-1.0	1.8E-04

Annexe 3 :

Tableau récapitulatif des OTUs différentiellement abondantes entre les stades de développement du riz dans la tige à l'issue du test des rangs signés de Wilcoxon sur les abondances relatives des OTUs chaque date de prélèvements

Abondance relative moyenne floraison 2017	Abondance relative moyenne maturation 2017	Abondance relative moyenne tallage 2018	Abondance relative moyenne floraison 2018	Abondance relative moyenne maturation 2018	Phylum	Class	Order	Family	Genus	median_diff	mean_diff	wilcox_p_value
10271.9	10724.5	10777.9	7993.1	12047.6	Proteobacteria	Alphaproteobacteria	Rickettsiales	Mitochondria	NA	2169.0	2731.4	1.43E-17
2953.9	2642.5	2638.1	3483.9	1218.6	Cyanobacteria	Oxyphotobacteria	Chloroplast	NA	NA	-987.0	-841.4	2.94E-07
0.0	1.8	0.6	162.4	2.1	Epsilonbacteraeota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	-2.5	-160.7	1.59E-06
0.1	0.9	0.2	37.3	1.3	Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Niveispirillum	-9.0	-36.4	5.98E-08
0.0	5.9	0.6	31.8	0.5	Bacteroidetes	Bacteroidia	Bacteroidales	Paludibacteraceae	Paludibacter	-0.5	-25.9	0.002
0.0	0.7	0.0	38.4	0.3	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Spirochaeta_2	-0.5	-37.7	7.33E-06
0.2	0.3	0.9	55.1	5.1	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	-11.0	-54.8	1.86E-10
0.0	0.8	1.7	50.4	2.0	Bacteroidetes	Bacteroidia	Flavobacteriales	Weeksellaceae	Chryseobacterium	-1.0	-49.6	2.54E-07
1.5	0.8	1.5	0.2	0.9	NA	NA	NA	NA	NA	-1.0	-0.7	0.0007
1.1	0.6	0.7	0.2	0.3	NA	NA	NA	NA	NA	-1.0	-0.5	0.0166

Annexe 4 :

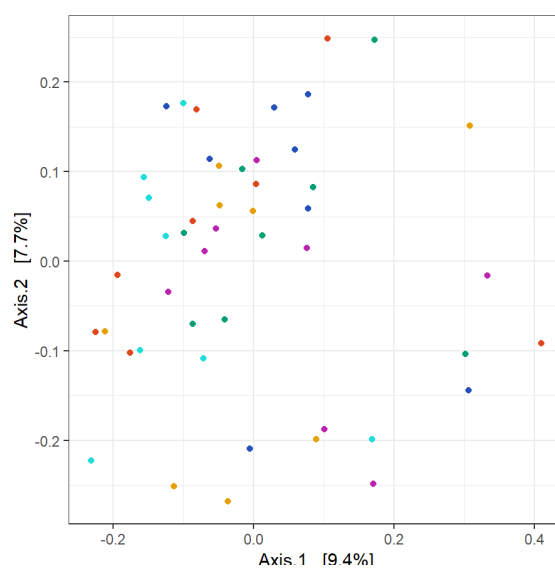
Résultats des analyses en coordonnées principales (PCoA) sur la distance UniFrac non pondérée pour tester l'effet du génotype de riz sur la structuration de ses communautés bactériennes endophytes

Variété de riz:

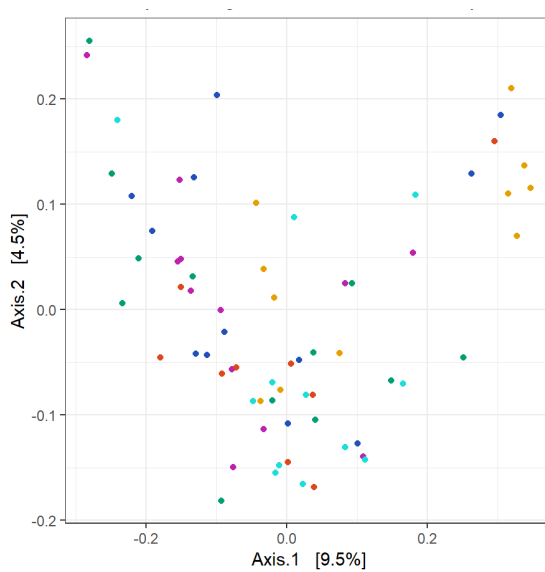
● Brio
 ● Gageron
 ● Mambo
 ● Paco
 ● RL 000X
 ● RV 000X

Racines de riz

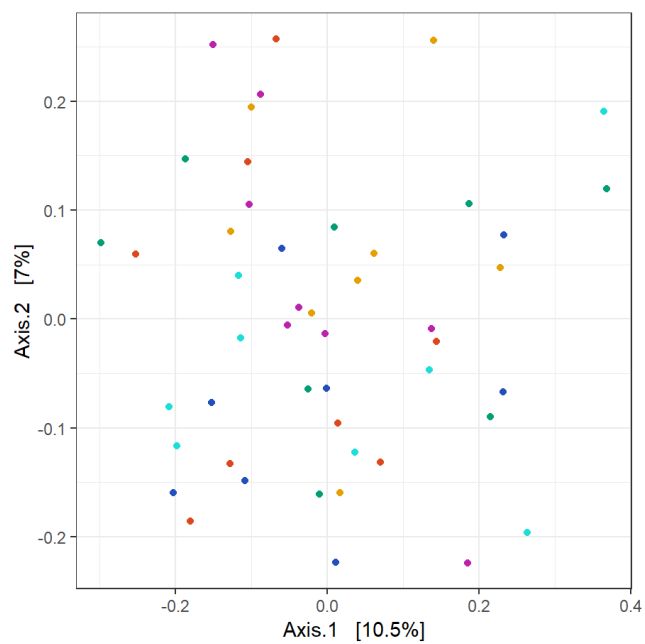
Stade tallage ; prélèvement 2018



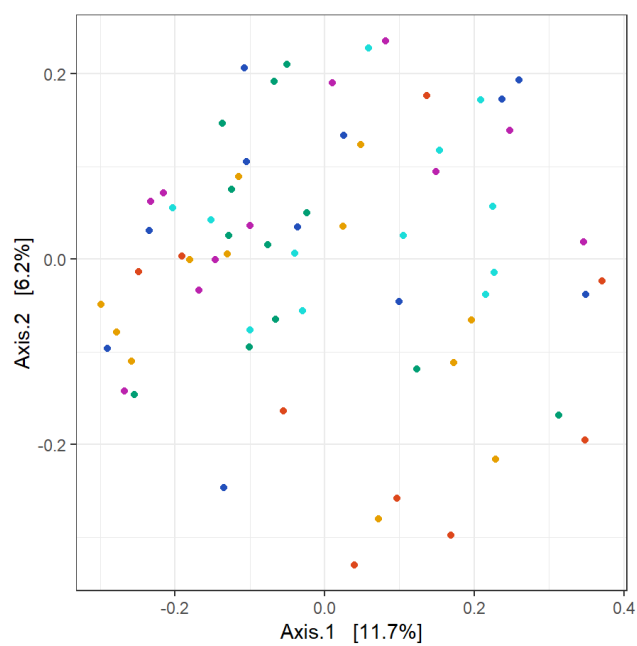
Stade Floraison ; prélèvement 2017



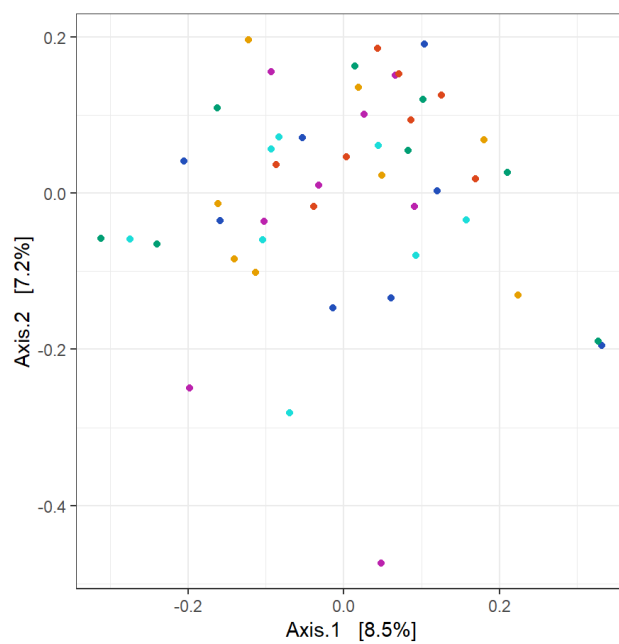
Stade Floraison ; prélèvement 2018



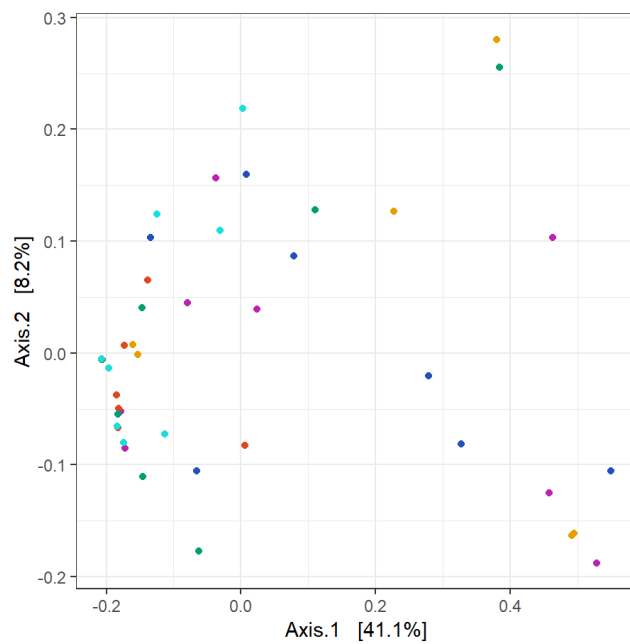
Stade maturation des grains ; prélèvement 2017



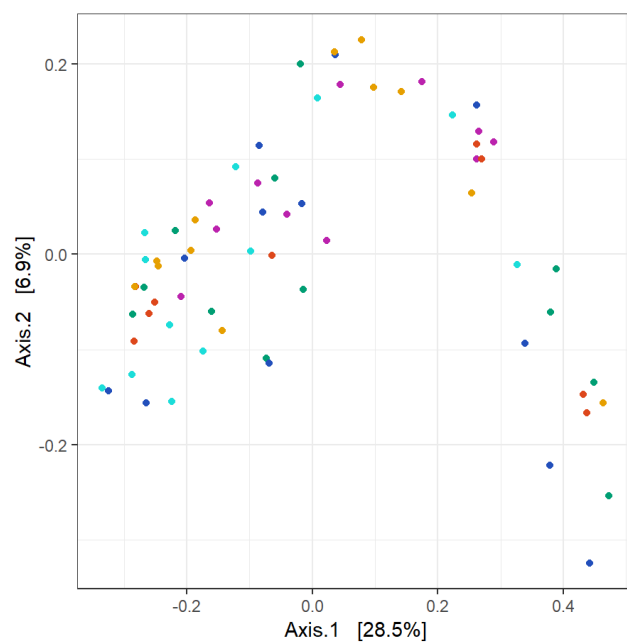
Stade maturation des grains ; prélèvement 2018

**Tiges de riz**

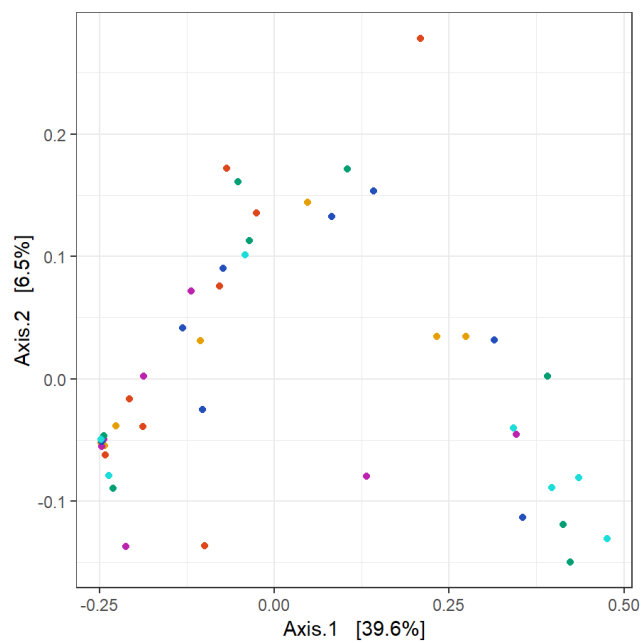
Stade tallage ; prélèvement 2018



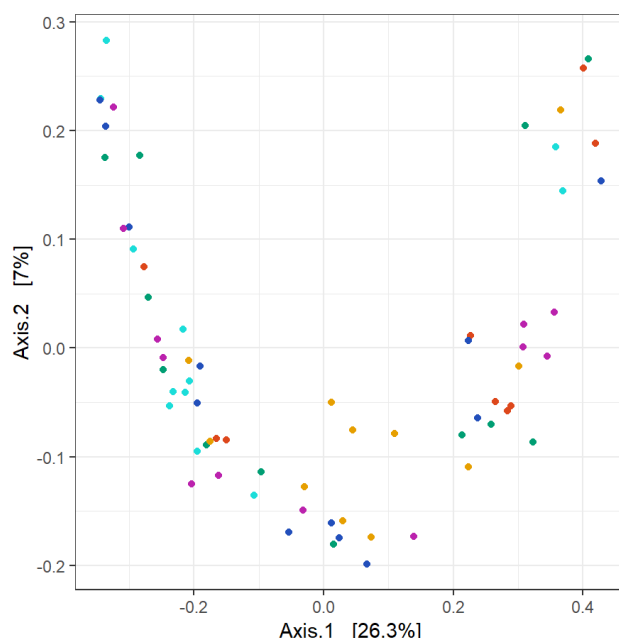
Stade Floraison ; prélèvement 2017



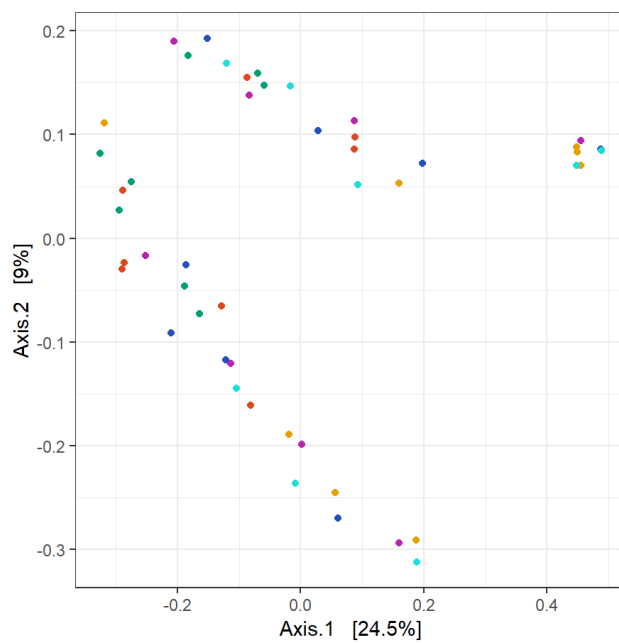
Stade Floraison ; prélèvement 2018



Stade maturation des grains ; prélèvement 2017



Stade maturation des grains ; prélèvement 2018



Annexe 5 :

Liste des OTUs fortement conservés (prévalence supérieure à 80%) et considérés comme appartenant au microbiote cœur des racines de riz pour chaque stade de développement et années de prélèvements, les croix indiquent pour chaque prélèvement la présence des OTUs avec une prévalence supérieure à 80%.

Phylum	Class	Order	Family	Genus	Species	all data set	I 2018	II 2017	II 2018	III 2017	III I2018
Campilobacterota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	na	x	x	x	x	x	x
Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Azospirillum	na	x	x	x	x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Azospira	na	x	x	x	x		x
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	na	x	x	x	x		x
Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Niveispirillum	na	x	x	x	x		
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	Rhizobium_ipomoeae	x	x	x	x		
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Hydrogenophaga	na		x	x	x		
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax	na		x		x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Propionivibrio	na			x	x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	na	na			x			x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	na	na				x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax	na		x	x			
Desulfobacterota	Desulfovibrionia	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	na				x		x
Verrucomicrobiota	Verrucomicrobiae	Opitutales	Opitutaceae	Lacunisphaera	na			x			x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Candidatus_Accumulibacter	na				x		x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Propionivibrio	na		x				x
Spirochaetota	Spirochaetia	Spirochaetales	Spirochaetaceae	Spirochaeta_2	Spirochaeta_aurantia					x	x
Campilobacterota	Campylobacteria	Campylobacterales	Sulfurimonadaceae	Sulfuricurvum	na			x	x		
Firmicutes	Clostridia	Peptostreptococcales	Acidaminobacteraceae	Acidaminobacter	na		x				
Proteobacteria	Alphaproteobacteria	Rhizobiales	Ancalomicrobiaceae	Ancalomicrobium	na				x		
Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Niveispirillum	na				x		
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Hydrogenophaga	na		x				
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Ideonella	Ideonella_sp.			x			
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	na	na						x
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	na	na			x			
Firmicutes	Clostridia	Lachnospirales	Lachnospiraceae	na	na				x		
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Magnetospirillaceae	Magnetospirillum	na				x		
Actinobacteriota	Actinobacteria	Corynebacteriales	Mycobacteriaceae	Mycobacterium	CNJ881_PLna4						x

Proteobacteria	Gammaproteobacteria	na	na	na	na	x			
Proteobacteria	Gammaproteobacteria	na	na	na	na	x			
Verrucomicrobiota	Verrucomicrobiae	Opitutales	Opitutaceae	Lacunisphaera	na				x
Bacteroidota	Bacteroidia	Bacteroidales	Paludibacteraceae	Paludibacter	na		x		
Bacteroidota	Bacteroidia	Bacteroidales	Paludibacteraceae	na	na				x
Proteobacteria	Alphaproteobacteria	Rhizobiales	Pleomorphomonadaceae	na	Kaistia_sp.		x		
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	na	x			
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	na	x			
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Propionivibrio	na		x		
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Sulfuritalea	na		x		
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Azovibrio	na		x		
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Dechloromonas	na	x			
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Novispirillum	Novispirillum_ittersonii		x		
Spirochaetota	Spirochaetia	Spirochaetales	Spirochaetaceae	Salinispira	na		x		
Spirochaetota	Spirochaetia	Spirochaetales	Spirochaetaceae	Treponema	na				x
Campilobacterota	Campylobacteria	Campylobacterales	Sulfurospirillaceae	Sulfurospirillum	na		x		