Université du Québec

INRS Institut national de la recherche scientifique

Centre Armand-Frappier Santé Biotechnologie

MODULATION DES COMMUNAUTÉS BACTÉRIENNES ET FONGIQUES ASSOCIÉES À UN HÔTE SOUS L'INFLUENCE DES STRESS BIOTIQUES ET ABIOTIQUES : CAS DE L'AGRILE DU FRÊNE

Par

Judith Jocelyne Mogouong Tambue

Thèse présentée pour l'obtention

du grade de *Philosophiæ* Doctor (Ph. D.)

en Biologie

Jury d'évaluation

Président du jury et Examinateur interne	Pr Charles Dozois INRS-Centre Armand-Frappier Santé Biotechnologie
Examinateur externe	Dr Armand Seguin Ressources Naturelles Canada
Examinateur externe	Dr Richard Trudel Consultant en Entomologie RT enr.
Directeur de recherche	Pr Claude Guertin INRS-Centre Armand Frappier Santé Biotechnologie
Codirecteur de recherche	Pr Philippe Constant INRS-Centre Armand Frappier Santé Biotechnologie

REMERCIEMENTS

Ces travaux de thèse n'ont été possible que par l'apport et le soutien moral et scientifique de plusieurs personnes.

Mes remerciements vont tout d'abord au Professeur Claude Guertin qui dès le début a cru en mes capacités, et dont les conseils et orientations pertinents ont permis de faire de ce travail à la fois une chance pour moi et une aventure passionnante. Malgré mon parcours qui n'a pas toujours été facile, sa posture rassurante m'a toujours aidé à canaliser mon énergie. MERCI Claude !

Merci au Professeur Philippe Constant qui m'a fait l'honneur accepter sans hésitation de m'accompagner dans cette aventure. Ses expertises très variées, son humour, sa patience infinie et surtout son optimisme sans faille m'ont permis de toujours y croire. Je me suis toujours sentie chanceuse de pouvoir bâtir ce projet avec mes deux directeurs de thèse.

Je me dois de remercier le Professeur Pierre Legendre pour son écoute et sa disponibilité. J'ai beaucoup apprécié les échanges dynamiques que nous avons eu pendant mon doctorat et surtout l'ouverture d'esprit de Pierre face à mes questionnements parfois déstructurés.

En m'engageant dans ce programme, j'appréhendais beaucoup d'être jugée pour mes choix, étant donné mon profil atypique… Je veux remercier Audrey-Anne Durand et Narin Srei pour m'avoir accepté dès le premier jour avec beaucoup d'empathie, dans une ambiance chaleureuse. Je ne saurais oublier les nombreuses discussions enrichissantes dans une bienveillance permanente devant nos paillasses qui m'ont à chaque fois permis d'ajuster mes approches. Cette ambiance a énormément beaucoup contribué à ma motivation. Sans le savoir, elles ont été pour beaucoup dans la réussite de ce long parcours.

Je remercie Professeur Charles Dozois, Dr Armand Séguin, et Dr Richard Trudel qui m'ont fait l'honneur d'accepter d'être les examinateurs de ma thèse. Le temps qu'ils ont bien voulu accorder à la relecture de ce manuscrit m'a permis de rectifier la forme, d'améliorer la clarté de plusieurs parties du texte et de compléter utilement certains points.

Je tiens également à remercier Josianne Guimond-Normandeau, Livie Lestin, Rose Ragot, Thibaut Bourdin, Adrien Frémont, ainsi que tous ceux (professeurs et étudiants) qui ont travaillé avec moi pour leur disponibilité et leur immense gentillesse. J'ai beaucoup apprécié les moments d'échanges stimulants partagés dans le fameux bureau étudiant avant l'arrivée du fameux Covid-19. Je n'oublie pas Michel Courcelles qui a toujours su m'aider à trouver réponses à mes questions lorsque j'étais perdue dans les méandres de mes bibliographies.

Un dernier mot et certainement le plus important est pour ma famille qui m'a soutenue sans faille pendant ces quatre années intenses. Mon mari et amoureux Christian Fioriti sans qui je n'aurai jamais pu aller au terme du programme. Il a su m'écouter, accepter mes humeurs instables et surtout parfois accepter que je préfère des échanges privés avec mon écran d'ordinateur à la place d'un bon moment ensemble. Je remercie aussi mes enfants Cendrine, Jérémy et Thomas qui ont accepté et encouragé les choix de leur maman dès le début de l'aventure. Leurs encouragements sans limite ont contribué à doper ma motivation à chaque fois que je m'interrogeais sur le sens de mes choix. Merci particulier à Thomas Fioriti qui a contribué à sa manière à mes travaux grâces à son écoute, sa contribution dans la réalisation de mes figures et la relecture de certaines notions. Enfin merci à mes parents qui m'ont transmis cette soif d'apprendre qui m'a animé tout au long de mon parcours. Nul doute que mon papa, Samuel Tambue, de là-haut est fière de me voir réaliser ce rêve que je caresse depuis mon jeune âge.

I believe the idea of species itself requires symbiosis. Lynn Margulis

AVANT PROPOS

Les trois premiers chapitres de la thèse sont présentés sous la forme d'articles scientifiques. Le premier publié dans *FEMS Microbiology Ecology* et le second publié dans *Scientific Reports.* Ces deux articles rapportent les résultats relatifs à la dynamique des communautés bactériennes et fongiques associées au tract intestinal de l'agrile du frêne confronté à des caractéristiques de son environnement, en se basant sur une approche de séquençage à haut débit d'amplicons de PCR. L'agrile du frêne, *Agrilus planipennis* Fairmaire, est un insecte ravageur s'attaquant spécifiquement à toutes les essences de frêne (*Fraxinus spp*., famille de *Oleaceae*). Il est considéré comme une espèce exotique invasive, responsable d'importants dommages écologiques et économiques en Amérique du Nord. Malgré l'existence de stratégies de lutte actuelle pour tenter de maitriser la propagation de l'agrile du frêne en Amérique du Nord, il reste très difficile de détecter des infestations initiales dans un délai raisonnable. Il existe très peu de connaissances disponibles à ce jour sur la contribution potentielle du microbiome intestinal des insectes ravageurs dans leur mécanisme d'adaptation dans son nouvel environnement. Les travaux de cette thèse visent à explorer la modulation des communautés bactériennes et fongiques associées à un hôte sous l'influence de stress environnementaux. Le premier article a permis de mettre en évidence un effet significatif de la taille de la population d'insectes au sein d'un arbre hôte sur la structure taxonomique de la communauté bactérienne associée au tract intestinal de l'insecte. L'analyse de l'alphadiversité et de la bétadiversité a permis de mettre en évidence l'effet discriminant de la taille de la population d'insectes au sein de l'arbre hôte pour la communauté bactérienne contrairement à la communauté fongique. Le deuxième article quant à lui a révélé, grâce à une analyse du partitionnement de la variation, une contribution significative du microbiome de la phyllosphère des arbres hôtes sur la structure taxonomique des communautés bactériennes et fongiques associées à l'insecte. Ce résultat suggère l'existence de relations complexes microbes-microbes au sein de chaque biotope (tract intestinal de l'insecte et phyllosphère de l'arbre hôte). Finalement, l'article 3, à être soumis, présente les résultats d'une analyse exploratoire de la complexité des communautés bactériennes et fongiques associées au tract intestinal de l'agrile du frêne. Il présente une analyse comparative de la structure phylogénétique et des réseaux de cooccurrences des microbiomes au sein leurs hôtes respectifs. L'ensemble de ces résultats sont repris dans une discussion générale permettant de mettre en évidence les résultats de cette thèse qui ouvrent de nouvelles avenues de recherche qui pourraient servir dans le cadre d'une approche de lutte biologique contre des insectes ravageurs.

RÉSUMÉ

« La vie ne resterait pas longtemps possible en l'absence de microbes » *Louis Pasteur*

L'agrile du frêne, *Agrilus planipennis* Fairmaire, est un insecte ravageur exotique originaire du continent asiatique qui cause d'importants dommages à toutes les espèces de frêne (*Fraxinus spp*.) en Amérique du Nord. Ce coléoptère de la famille des Buprestidae est considéré, aujourd'hui, comme étant le plus sérieux ravageur forestier des dernières décennies, causant d'importants impacts environnementaux, écologiques et économiques. Les communautés microbiennes associées au tract intestinal des espèces d'insectes exotiques envahisseurs pourraient jouer un rôle crucial dans leurs capacités d'adaptation à leur nouvel environnement. À l'inverse, des facteurs environnementaux tels que la température et la nutrition pourraient agir comme déterminants et influencer la composition du microbiome de ces insectes. L'objectif de cette thèse est d'explorer les modulations du microbiome de l'agrile du frêne face aux variables environnementales, afin de mettre en évidence la dynamique du microbiome de cet insecte ravageur confronté à un nouvel environnement. L'approche choisie pour l'ensemble des travaux s'est appuyée sur la technique de séquençage d'amplicon ciblant le gène codant de l'ARNr 16S bactérien et de l'espaceur transcrit interne fongique ITS2 et une variété d'analyses multivariées. Dans un premier temps, une relation significative a été identifiée entre le nombre de captures des adultes de l'agrile du frêne sur des arbres hôtes et la structure taxonomique de la communauté bactérienne associée au tract intestinal de l'insecte. En effet, en analysant la variation des indices de diversité et les distances de dissimilarité de ces communautés, il a été possible de démontrer que la densité de la population d'insectes affectait leurs structures. De plus, certains taxa bactériens se sont révélés comme étant des indicateurs spécifiques pour les insectes capturés sur des arbres à faible densité de population. Dans un autre volet de l'étude, on cherchait à déterminer, par une analyse du partitionnement de la variation, les contributions relatives des propriétés spécifiques des arbres hôtes, à savoir les communautés bactériennes et fongiques associées aux feuilles (phyllosphère), la composition phytochimique de ces dernières, et les distances géographiques des sites d'échantillonnage sur la communauté microbienne intestinale de l'agrile du frêne. La composition du microbiome de la phyllosphère s'est révélée être un puissant prédicteur de la structure de la communauté microbienne dans les intestins de l'agrile du frêne, expliquant respectivement 53 et 48 % de la variation des champignons et des bactéries. Une comparaison des deux biotopes (tract intestinal de l'insecte et phyllosphère) a révélé des

structures microbiennes taxonomiques distinctes avec des taxons indicateurs spécifiques pour chacun. Finalement, on a cherché à comprendre la complexité des relations existant dans les deux biotopes, en ciblant une approche exploratoire comparative des structures et réseaux de cooccurrences des communautés microbiennes. La tendance au regroupement phylogénétique des espèces bactériennes et fongiques a mis en évidence le rôle déterminant du biotope dans l'assemblage des espèces coexistant dans chaque habitat, alors que le réseau de cooccurrences est apparu avec une densité plus élevée pour les espèces cooccurrentes dans le tract intestinal de l'insecte. Cette étude, basée sur des méthodes novatrices, a permis de proposer une nouvelle approche plus adaptée pour les travaux de recherche relatifs à la plasticité du microbiome, qui permet une interprétation plus adéquate des résultats d'analyses. Par exemple, nos résultats ont suggéré des stratégies de modulation complexes du microbiome intestinal des insectes envahissants confrontés à certains facteurs environnementaux ouvrant de nouvelles perspectives de recherche.

Mots-clés : *Agrilus planipennis* Fairmaire; Agrile du frêne; Diversité microbienne; Perturbation; Microbiome intestinal; Taille de la population d'insectes; Communauté microbienne; *Fraxinus spp*.; Frêne; Interactions microbe-plante-insecte; Phyllosphère; Partitionnement de la variation.

ABSTRACT

“Life would not long remain possible in the absence of microbes” *Louis Pasteur*

The Emerald Ash Borer, *Agrilus planipennis* Fairmaire, is an exotic insect pest native to Asia that causes environmental and economic damage to ash trees (*Fraxinus spp.)* in North America. This Buprestidae first detected in 2002 at Canada-U.S. border, has been reported as the most serious pest in recent decades, particularly because of the environmental, ecological and economic impacts. As microorganisms are reported to contribute vital functions for their hosts, the microbial communities associated with the gut of invasive alien insect species could play a crucial role in their ability to adapt to their new environment. On the other hand, environmental factors such as temperature and nutrition could act as determinants, and influence the composition of these insects’ microbiome. The objective of this thesis is to explore the modulations of the microbiome of the emerald ash borer when facing some environmental variables in order to highlight the potential role of the insects’ microbiome during its establishment process in North America. The overall approach was based on the amplicon sequencing high throughput technique targeting the bacterial 16S rRNA and the ITS2 fungal internal transcribed spacer couple with different multivariate analyzes. First, a significant relationship was found between the insect’s population size in the host trees and the taxonomic structure of bacterial community associated with the insect’s gut. In fact, by analyzing the variation in the diversity indices and the dissimilarity distances of the bacterial communities, it was possible to demonstrate that the taxonomic structure of microbial communities was affected. Moreover, some bacterial taxa appeared as indicator taxa for the lowest insect population size. Secondly, the insects’ gut variation and its partitioning performed after some canonical analyzes revealed significant contributions of host tree-specific properties. More specifically, the bacterial and fungal communities associated with the host leaves (phyllosphere) appeared to play a determinant role in the variation observed in the insects’ microbiome compared to geographic coordinates and some leaves’ phytochemicals. The composition of the phyllosphere microbiome has been shown to be a powerful predictor of the structure of the microbial community in the intestines of the emerald ash borer, accounting for 53 and 48% of the variation in fungi and bacteria, respectively. A comparison of the two biotopes (insects’ gut and phyllosphere) revealed distinct taxonomic microbial structures with specific indicator taxa for each of them. Finally, an attempt was made to understand the complexity of the

relationships in both biotopes, by targeting a comparative approach based on phylogenetic structures and cooccurrence networks of the microbial communities. Thus, the biotopes were found playing a determinant role in the assembly of the microorganisms in each biotope due to the tendency of the cooccurring species to be phylogenetically clustered whereas the cooccurrences network appeared with a higher density in the insect's gut. This study using innovative approaches pointed out some modulation strategies related to the insects' gut microbiome of an invasive insect facing some environmental factors that should be considered in the future for potential biological pest management.

Keywords : *Agrilus planipennis* Fairmaire; emerald ash borer; microbial diversity; disturbance; gut microbiome; insect population size; microbial community; *Fraxinus spp.*; ash; microbe-plants-insect interactions; phyllosphere; variation partitioning.

TABLE DES MATIÈRES

LISTE DES FIGURES

LISTE DES TABLEAUX

LISTE DES ÉQUATIONS

$\{0, N_1 + N_2 - N\} \leq j \leq min\{N_1, N_2\}$:

$P_j = \frac{C(N,j)\ x\ C(N-j,N_2-j)\ x\ C(N-N_2,N_1-j)}{C(N,N_2)\ x\ C(N,N_1)}$

$P_{lt} = \sum P_j\ pour\ j = 0$ à $Q_{Obs} - 1$

$P_{gt} = \sum P_j\ pour\ j = Q_{Obs} + 1$ à N

Q_{Obs} : la fréquence des cooccurrences observées

N : nombre total de sites

P_j : la probabilité que deux espèces apparaissent exactement à un site spécifique j (les valeurs de j varient de 0 à N).

P_{lt} : la probabilité que deux espèces apparaissent à une fréquence plus faible que la fréquence d'apparition observée

P_{gt} : la probabilité que deux espèces apparaissent à une fréquence plus élevée que la fréquence d'apparition observée

$Q = \frac{1}{2m}\sum_{uw}[A_{uw} - \frac{k_u k_w}{2m}]\ \delta(c_u, c_w)$ = Modularité

m : nombre d'interactions dans le graphe

A_{uw} : un élément de la matrice adjacente du réseau d'interactions, les valeurs varient de 0 (pas de connexion entre u et w) et 1 (connexion entre u et w)

k_u / k_w : degré du nœud u ou w correspond au nombre de connexions qui leur sont attachées

c_v / c_w : point de connexion (nœud) u ou w appartenant à la communauté

LISTE DES ABRÉVIATIONS

ADN : Acide Deoxyribonucléique

ANOVA : *Analysis of variance* (Analyse de la variance)

ARNr : Acide ribonucléique ribosomal

ASV : *Amplicon sequence variant* (Variants de séquences d'amplicons)

CCi : *Closeness centrality* (Centralité de proximité)

C/N : ratio Carbone/Azote

EAB : Emerald Ash Borer (Agrile du frêne)

MPD : *Mean pairwise distance* (distance moyenne par paire)

MNTD : *Mean nearest taxon distance* (distance du taxon le plus proche)

OTU : *Operational taxonomic unit* (Unité taxonomique opérationnel)

PCA : *Principal componant analysis* (Analyse en composantes principales)

PCoA : *Principal coordinate analysis* (Analyse en coordonnées principales)

PCR : *Polymerase chain reaction* (Réaction polymérase en chaine)

PD : *Phylogenetic Diversity* (Diversité phylogénétique)

PERMANOVA : *Permutational multivariate analysis of variance* (Analyse multivariée par permutation de la variance)

RDA : *Redondancy analysis* (Analyse de redondance)

SIMPROF : *Similiarity Profile Analysis* (Analyse de profile de similarité)

UPGMA : *Unweighted pair group method with arithmetic mean* (Méthode de groupe de paires non pondérées avec moyenne arithmétique)

INTRODUCTION

1.1 Agrile du frêne

L'agrile du frêne, *Agrilus planipennis* Fairmaire, est un insecte ravageur s'attaquant spécifiquement à toutes les essences de frêne (*Fraxinus spp*., famille de *Oleaceae*). Il est originaire de l'Asie (Chine, Corée, Japon, Taiwan, Laos, Mongolie) et de l'Est de la Russie (Jendek & Grebennikov, 2011; Wei *et al.*, 2007). Cet insecte de l'ordre des coléoptères et de la famille des Buprestidae correspond à la 8^{e} plus grande famille des scarabées au monde avec environ 14,700 espèces (Bellamy, 2002). Il appartient au genre *Agrilus spp*. avec plus de 1,000 espèces dont certaines sont considérées comme étant de sérieux ravageurs, responsables de la mortalité d'arbres dans plusieurs forêts.

1.1.1 Biologie de l'insecte

Les adultes de l'agrile du frêne sont caractérisés par des dimensions assez variables (la longueur varie entre 8,8 et 14 mm, largeur de 3,1 à 3,4 mm) et une couleur vert métallique caractéristique (Valenta *et al.*, 2017; Wei *et al.*, 2007). Des caractéristiques morphologiques permettent de différencier les mâles et les femelles. En effet, les mâles sont généralement de plus petites tailles, alors que l'abdomen des femelles est plus large et légèrement bombé (Wang *et al.*, 2010). Généralement, les femelles pondent leurs œufs dans les crevasses de l'écorce. Après une incubation d'environ deux semaines, l'éclosion des œufs permet aux larves d'émerger, pour ensuite s'infiltrer sous l'écorce où elles se nourrissent sur le tissu cambial de l'intérieur de l'écorce et de l'extérieur du bois. Pendant leur nutrition et leur développement, les larves construisent des tunnels en zigzag en dessous de l'écorce. Une fois matures, elles creusent un tunnel dans le bois, construisant ainsi la chambre dans laquelle elles se transforment en pupe. Après maturation, la pupe donne un insecte adulte qui émerge quelques semaines plus tard en creusant un tunnel vers la surface. Le trou de sortie se présente sous forme de « D » sur l'écorce. La durée du cycle de vie de l'agrile du frêne (Figure 1.1) s'étend sur une période d'un à deux ans. En effet, la durée peut varier en fonction de plusieurs paramètres environnementaux (Wu, 2016; Wylie & Speight, 2012), notamment la température (plus la température est élevée, plus la durée des stades va être courte), la latitude (plus on va vers le nord, plus le cycle est long (Lyons *et al.*, 2007)), la densité des larves sous l'écorce et l'état de stress des arbres hôtes (Cappaert *et al.*, 2005b; Herms & McCullough, 2014; Orlova-Bienkowskaja & Bieńkowski, 2016; Wei *et al.*, 2007).

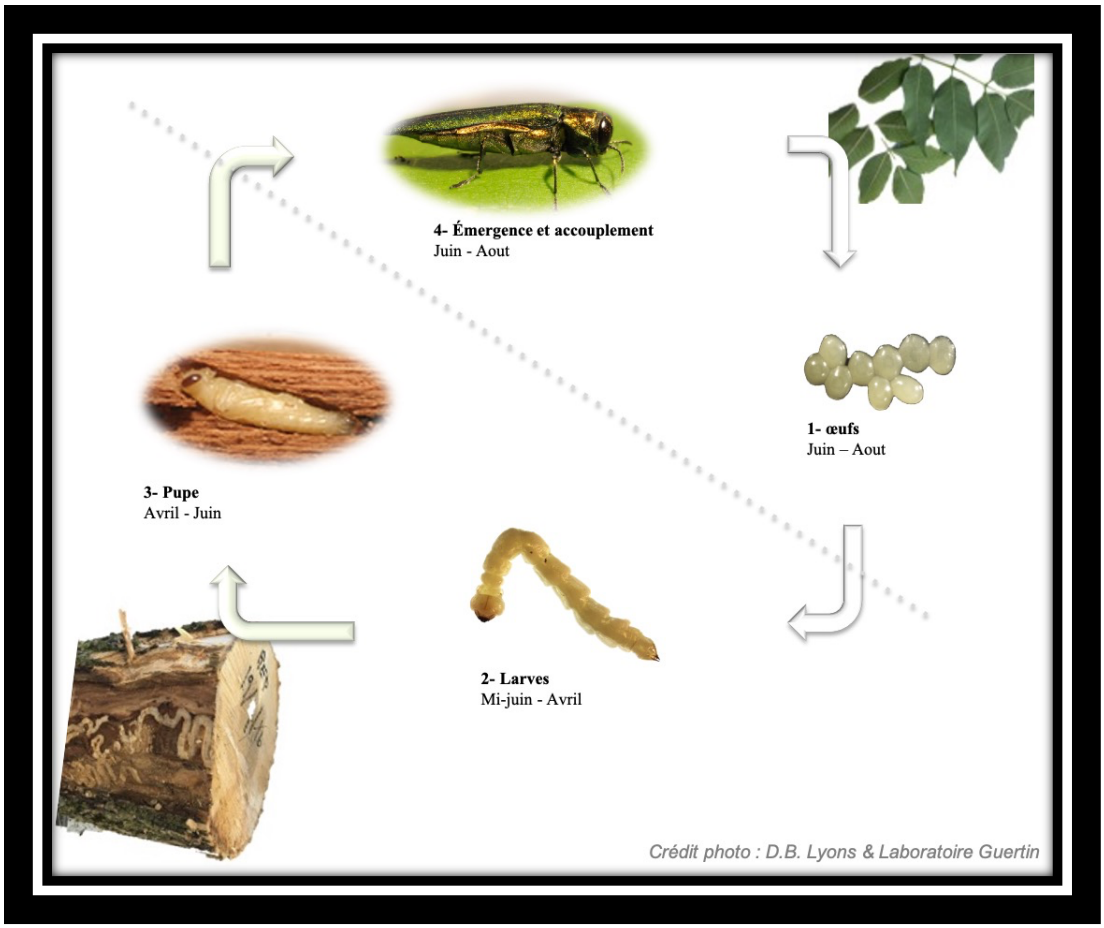


Figure 0.1 **Représentation du cycle de vie de l'agrile du frêne, Photographies laboratoire Guertin. Les durées indiquées sont basées sur les données disponibles relatives au développement de l'insecte en Amérique du Nord.**

Parmi les quatre étapes du cycle de développement, le stade larvaire est l'étape qui s'étend sur une durée plus longue (**Figure 0.1**). En Amérique du Nord, ce stade larvaire se fait en quatre principaux stades et s'étale de la fin de l'été jusqu'au printemps de l'année suivante (la larve hiberne à partir du quatrième stade) (Cappaert *et al.*, 2005b). Lorsque les larves se développent sous l'écorce, elles se nourrissent dans le tissu cambial empêchant ainsi la circulation de l'eau et des nutriments vers l'ensemble des parties de l'arbre. Au fur et à mesure de l'augmentation de la densité des larves sous l'écorce, les signes d'infestation apparaissent progressivement au niveau du feuillage et de la canopée de l'arbre. Des données collectées ont permis de montrer qu'environ 80-120 agriles adultes pouvaient se développer par m^2 avant que l'arbre ne succombe (McCullough *et al.*, 2007). Dès l'apparition des signes visibles au niveau de la canopée, l'arbre meurt dans les 4 à 6 années suivantes en Amérique du Nord (Herms & McCullough, 2014; Knight *et al.*, 2013; Poland & McCullough, 2006). Une étude a déterminé que le dépérissement de la

canopée devient apparent pour les frênes blancs et verts quand la densité de l'émergence atteint 25-35 adultes par mètre carré (Anulewicz, 2006).

L'agrile du frêne peut être considéré comme un insecte oligophage puisqu'il correspond à la définition de Harborne (1988) comme « un insecte se nourrissant sur un nombre d'espèces de plantes relativement réduit ». Plusieurs études s'accordent aujourd'hui sur le fait que l'insecte serait guidé vers la plante par la qualité nutritionnelle et les types de composés chimiques qui sont émis par l'arbre (Harborne, 1998). Alors que dans son habitat d'origine, l'agrile du frêne s'attaque principalement aux arbres stressés, en Amérique du Nord il est en mesure de s'attaquer aussi aux frênes en bonne santé (Cappaert *et al.*, 2005b; Haack *et al.*, 2002; Herms & McCullough, 2014; Poland & McCullough, 2006; Wei *et al.*, 2007). Une étude a suggéré d'une part que les adultes de l'agrile du frêne préfèreraient les feuilles exposées au soleil, et d'autre part que cette préférence serait aussi guidée par la concentration en protéines totales des feuilles (Chen & Poland, 2009). Sur un tout autre sujet, la nature des attaques de l'agrile du frêne, notamment la période cryptique du développement larvaire qui peut s'étaler jusqu'à deux ans, rend la détection d'une attaque initiale très difficile (Ryall *et al.*, 2011). Lorsque les signes d'infestation apparaissent sur les arbres, il est souvent trop tard pour lutter contre l'insecte (Lyons *et al.*, 2007). La **Figure 0.2** représente quelques signes d'infestation visibles sur un frêne attaqué.

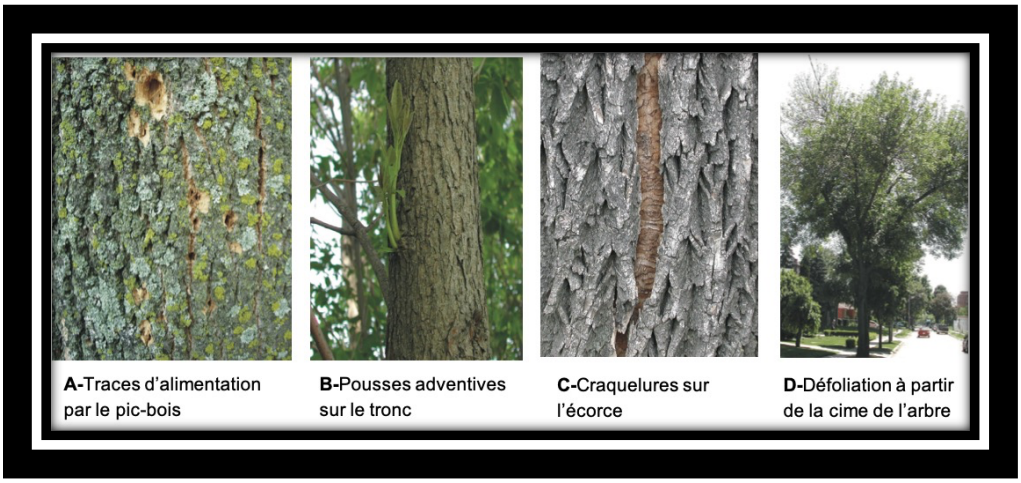


Figure 0.2 **Quelques représentations de signes d'infestation (Lyons *et al.*, 2007)**

1.1.2 Historique de l'émergence et répartition géographique de l'insecte

Natif de la région Asie de l'est, l'agrile du frêne a connu une progression fulgurante depuis sa première détection en 2002 à la frontière canado-américaine. Siegert *et al.* (2014) ont rapporté dans leurs travaux basés des données de dendrochronologie que l'agrile serait arrivé dans les années 1990 dans les banlieues de la ville de Détroit (État du Michigan). En janvier 2021, la présence de l'agrile du frêne a été confirmée dans cinq provinces canadiennes et 35 états américains. Le **Tableau 0.1** récapitule la distribution de l'insecte en date d'aujourd'hui. Par ailleurs, plusieurs études font état de la progression de l'insecte en Europe à partir de la Russie (Baranchikov *et al.*, 2008; Orlova-Bienkowskaja, 2014). En effet, plusieurs études ont suggéré que l'agrile du frêne était capable de se nourrir sur des frênes natifs d'Europe (McCullough, 2020; Pureswaran & Poland, 2009).

Tableau 0.1 Tableau récapitulatif des lieux où la présence de l'agrile du frêne a été confirmée.

Pays	Année détection/origine	Invasif	Référence
Chine	2002/Natif	Non	Haack *et al.* (2002) ; Wei *et al.* (2004)
Corée du Sud	2002/Natif	Non	Haack *et al.* (2002)
Corée du Nord		Non	CABI, EPPO, 2006. Agrilus planipennis. [Distribution map]. In: Distribution Maps of Plant Pests, Wallingford, UK: CABI. Map 675.
Russie	2002/Natif	Non	Haack *et al.* (2002)
Canada (5 provinces) - Nouveau Brunswick, 2018 - Nouvelle Écosse, 2018 - Manitoba, 2017 - Québec 2008 - Ontario, 2002	2002/Introduit	Oui	CFIA (2020) ; Haack *et al.* (2002)
États-Unis-d'Amérique (35 États)	2002/Introduit	Oui	Francese *et al.* (2011) ; Haack *et al.* (2002) ; USDA (2009)
Russie centrale (Moscou, 2007)	2008/Introduit	Oui	Baranchikov *et al.* (2008) ; Orlova-Bienkowskaja (2014)
Ukraine	2020/Introduit	Oui	Orlova-Bienkowskaja *et al.* (2020)

1.1.3 **Dommages occasionnés par l'infestation et stratégies de gestion**

L'agrile du frêne est considéré comme une espèce invasive et la plupart des infestations induisent la mortalité des frênes dont la sensibilité varie selon l'espèce. Dans sa région native, les adultes de l'agrile du frêne sont connus pour se nourrir essentiellement sur deux espèces de frênes, soit *F. chinensis* Roxb. et *F. mandshurica* Rupr. (Liu *et al.*, 2004). En Amérique du Nord, dans son aire de distribution actuelle, ils se nourriraient sur les feuilles du frêne vert, *F. pennsylvanica* Marsh., le frêne blanc, *F. americana* L. et le frêne noir, *F. nigra* Marshall. De plus, ces hôtes privilégiés qui sont largement distribués en Amérique du Nord seraient beaucoup plus susceptibles d'être infestés par l'agrile du frêne en comparaison avec le frêne bleu, *F. quadrangulata* Michx, plus résistante (Anulewicz, 2006). Cette résistance pourrait être attribuée à la présence d'un composé phénolique spécifique se trouvant dans les feuilles, le pinorésinol dihexoside (dérivé de lignines) et à la présence dans son phloème de fortes concentrations de lignines, coumarines, proline, tyramine et protéines de défense (Chakraborty *et al.*, 2014; Villari *et al.*, 2016; Villari *et al.*, 2014). Les espèces de frênes les plus distribués en Amériques étant le frêne vert et le frêne blanc, les travaux de Burr and McCullough (2014) rapportaient un taux de mortalité pour le frêne vert d'environ 79% dans les forêts du Sud-Est du Michigan. Ainsi, la propagation continue de l'insecte représente une réelle menace pour les ressources en frênes en Amérique du Nord et, par conséquent, une menace pour une variété d'écosystèmes du fait de l'abondance de ces trois espèces dans divers habitats (MacFarlane & Meyer, 2005).

Depuis sa détection, l'agrile du frêne serait également responsable d'importants dommages écologiques et économiques. Toutes les essences de frêne présentes en Amérique du Nord peuvent être attaquées par cet insecte, et sa propagation cause de nombreux dommages. Au niveau écologique, on peut citer la réduction de la biodiversité avec une influence directe et indirecte sur les communautés natives (Gandhi & Herms, 2010a; Gandhi & Herms, 2010b), la perturbation des écosystèmes avec la perturbation des cycles des nutriments et la perturbation de l'habitat avec par exemple la propagation des plantes photosensibles (Herms & McCullough, 2014). Tandis qu'au niveau économique, l'agrile du frêne est aujourd'hui considéré comme étant l'insecte ravageur le plus couteux et le plus destructif identifié en Amérique du Nord à ce jour (Aukema *et al.*, 2011). Les coûts liés aux traitements des arbres infestés, au reboisement et aux contraintes relatives à la gestion des zones de règlementation établies par le gouvernement sont considérables. Il a été estimé qu'entre 2009 et 2019, environ 17 millions de frênes des zones urbaines à travers 25 États des États-Unis auront besoin d'un traitement, d'un enlèvement ou d'un remplacement pour un montant évalué à 10,7 milliards de dollars américains (Kovacs *et al.*,

2010). Les propriétés ayant des frênes infestés ont vu la valeur immobilière de leurs biens chuter (Herms & McCullough, 2014). De plus, des scientifiques estiment que les coûts reliés au traitement et à l'enlèvement des arbres infestés par l'agrile du frêne dans les municipalités pourraient atteindre deux milliards de dollars sur une période de 30 ans avec des conséquences considérables sur le plan écologique (CFIA, 2020).

Les stratégies de lutte envisagées à ce jour visent principalement la prévention permettant d'offrir aux frênes infestés 1 à 2 années supplémentaires de vie (McCullough, 2020). On distingue principalement la lutte biologique et l'utilisation d'insecticides. L'utilisation d'insecticides systémiques a permis de développer de multiples traitements. Le coût associé à leur utilisation serait plus faible que ceux associés à l'enlèvement des arbres infestés (Herms & McCullough, 2014). La lutte biologique envisagée par certaines municipalités de la province du Québec est basée soit sur l'utilisation de parasitoïdes, soit sur l'utilisation de champignons entomopathogènes. Parmi les espèces de parasitoïdes utilisées à ce jour, on peut citer le *Oobius agrili* Zhang and Huang (parasitoïdes pour les œufs d'agrile), *Tetrastichus planipennisi* Yang (parasitoïdes pour les larves d'agrile) et *Spathius agrili* Yang (parasitoïdes pour les larves) (Duan *et al.*, 2017). Le laboratoire Guertin travaille depuis plusieurs années sur le développement d'approches pouvant être intégrées dans des stratégies de contrôle des populations d'insectes envahisseurs dans une approche de lutte microbiologique. Ainsi, il a contribué à l'élaboration du dispositif d'autodissémination de conidies d'un isolat du champignon entomopathogène *Beauveria bassiana* (Balsalmo) Vuillemin permettant de contrôler les populations de l'agrile de frêne (Lyons *et al.*, 2012; Srei *et al.*, 2020a; Srei *et al.*, 2020b). L'utilisation des pièges est très répandue pour la capture d'insectes avec par exemple l'ajout de phéromones et/ou de kairomones pour optimiser les chances de capture. Parmi les deux types de pièges disponibles (pièges prismes et les pièges à entonnoirs (pièges Lindgren ayant 8 ou 12 entonnoirs successifs)), les pièges à entonnoirs présentent l'avantage d'être réutilisables, faciles à transporter et la récupération des insectes est immédiate et ces pièges sont significativement plus efficaces pour détecter des arbres nouvellement infestés (Francese *et al.*, 2011; Lyons *et al.*, 2012). Toutefois, à ce jour, il n'existe pas de technique reproductible permettant la détection d'insectes en début d'infestation. Des progrès doivent donc encore être faits pour développer des outils de détection plus performants capables d'indiquer les débuts d'attaque.

1.2 Attaque par les herbivores : facteurs de stress des plantes

Le mode d'alimentation herbivore caractérise environ 45% des espèces d'insectes et les interactions entre les insectes herbivores et les plantes sont bien documentées. D'après Marquis (2004) l'évolution des insectes pourrait avoir joué un rôle important dans le façonnage du monde végétal. Malgré le fait que les tissus végétaux sont difficilement digérés par les insectes, le mode d'alimentation herbivore est répandu chez les insectes en dépit de l'absence de certains enzymes permettant leur digestion. Bien que l'impact de l'action des insectes herbivores ne tue pas toujours les plantes, il peut avoir entrainé des changements au niveau des caractéristiques chimiques et physiologiques (Schoonhoven *et al.*, 2005; Wielkopolan & Obrepalska-Steplowska, 2016). En effet, l'herbivorie par les insectes est un des facteurs induisant un stress chez les plantes pouvant déclencher des mécanismes de défense. Les signes visuels et des indicateurs biochimiques peuvent permettre de caractériser le niveau de stress d'un arbre infesté par des insectes. Chez les dendroctones par exemple, Slansky and Haack (1986) ont montré que les insectes seraient affectés par des composés secondaires se trouvant dans les écorces, les feuilles ou le phloème lorsque ces derniers s'y nourrissent ou encore lorsqu'ils cherchent à s'accoupler. Dans le cas des insectes s'attaquant au phloème, les travaux de Eyles *et al.* (2010) suggèrent que le mécanisme de défense de l'arbre pourrait avoir trois composantes. Ainsi, la première serait la « préattaque » où des défenses physiques et des composés chimiques présents dans le phloème et potentiellement toxiques pour l'insecte sont mis à contribution, suivie par « l'attaque induite » où la synthèse de nouveaux métabolites secondaires à faible poids moléculaire, des polymères phénoliques (lignine et subérine) et des protéines protectrices sont produits et accumulés, et enfin la « formation de tissus » servant à protéger les blessures. Tandis que selon Wielkopolan and Obrepalska-Steplowska (2016), le mécanisme de défense se ferait en deux étapes : d'une part l'attaque directe entrainant l'inhibition de comportements importants chez l'insecte, tels que la reproduction et l'alimentation, en contribuant notamment à la protection mécanique de la surface des plantes avec un durcissement de la paroi cellulaire par le processus de lignification et d'autre part à la production de molécules toxiques pour les insectes. Par exemple, l'acide jasmonique qui est une molécule ubiquitaire présente dans les plantes jouerait plusieurs fonctions dans la défense contre les insectes ravageurs. Il agirait comme un stimulateur clé pour la production de substances volatiles par la plante. Les quantités d'acide jasmonique augmentent généralement lorsque les plantes sont soumises à des stress liés à des blessures mécaniques (Rodriguez-Saona *et al.*, 2006). Ce composé semble unanimement considéré comme le régulateur le plus important dans le mécanisme de défense des plantes face aux insectes (Wielkopolan & Obrepalska-Steplowska, 2016). Ainsi, les attaques conduiraient à

l'émission de l'acide jasmonique et de l'éthylène qui induisent deux types de réactions chez la plante : la synthèse de différentes protéines parmi lesquelles on compte les protéines anti-nutritionnelles ou protéines de défense (lécithines, inhibiteurs de protéases et inhibiteurs d'alpha-amylase), et la synthèse de molécules anti-insectes. Ainsi, cette étude rapporte une augmentation de la teneur en MeJA (methyl jasmonate, dérivé de l'acide jasmonique) dans l'écorce de deux espèces de frêne (*F. mandshurica* et *F. americana*) qui induirait une accumulation de la teneur en phénols dans le phloème de ces arbres. Les composés anti-insectes émis par des feuilles mâchées peuvent être de différentes natures : des terpénoïdes, des composés volatils des feuilles vertes et des composés azotés (Arimura *et al.*, 2009; Paré & Tumlinson, 1999; Wielkopolan & Obrepalska-Steplowska, 2016). Des activités variées ont été rapportées pour des composés phytochimiques extraits des frênes : les composés phénoliques ayant une activité anti-alimentaire et antibactériens (Chen *et al.*, 2011; Harborne, 1988; Kostova & Iossifova, 2007; Villari *et al.*, 2014), la lignine jouant un rôle dans la défense contre les pathogènes et les insectes, des substances organiques naturelles comme la coumarine ayant des activités antibactériennes et inhibitrices (Kostova & Iossifova, 2007). D'après les travaux de Villari *et al.* (2014), le stress induit par l'ajout de MeJA favorise des niveaux élevés en verbascoside et lignine chez les frênes *F. americana* et *F. pennsylvanica*. Enfin, ces composés peuvent être des toxines et des inhibiteurs digestifs (Villari *et al.*, 2014). Les travaux de Chen *et al.* (2011) ont suggéré une variation de la réponse dans la chimie foliaire de trois espèces de frêne (*F. nigra, F. pennsylvanica et F. americana*) soumis au stress de l'attaque par des insectes herbivores. En effet, d'après cette même étude il existe d'autres composés phytochimiques qui peuvent affecter les insectes. Ce sont des composés azotés de défense comme des inhibiteurs de chymotrypsine qui contribuent à la diminution de la valeur nutritive des feuilles, et les inhibiteurs de protéase qui sont produits par les arbres à la suite d'une attaque peuvent quant à eux inhiber la croissance des insectes. L'augmentation des quantités d'inhibiteurs de la chymotrypsine liée à l'alimentation des insectes a été mise en évidence chez *F. americana* (Chen *et al.*, 2011). De plus, il y a aussi les composés hydrocarbonés non structuraux totaux (TNC) que l'on retrouve dans les cellules des plantes et sont généralement plus digestes que les hydrates de carbone structuraux. C'est ainsi qu'il a été mis en évidence une augmentation des taux de TNC dans des conditions ensoleillées, lorsque l'arbre était blessé et dans les feuilles matures (Chen & Poland, 2009).

1.3 Microorganismes associés aux insectes

« Microorganisms are widely present in ecosystems, but usually as hidden players" (Schoonhoven et al., 2005).

1.3.1 Concept du microbiome

Les eucaryotes sont apparus sur la terre au moins deux millions d'années avant les bactéries et auraient évolué depuis leur apparition en présence de multiples communautés microbiennes (Margulis, 1993). Ainsi, les microorganismes tels que les bactéries, les champignons, les archées et les protozoaires colonisent les êtres vivants et sont désignés sous le terme microbiote (Feldhaar, 2011). Le terme microbiome sera utilisé dans ce projet pour représenter l'ensemble du génome de microbiote. En se basant sur la majorité des travaux de recherches faites jusqu'à ce jour, on sait maintenant que la majorité des microorganismes sont essentiels pour l'écosystème, notamment grâce à leurs interactions bénéfiques avec d'autres microorganismes environnants ou avec les organismes hôtes (Berg *et al.*, 2020). Ainsi, le microbiote serait capable d'aider les animaux à accomplir leurs fonctions physiologiques telles que les fonctions relatives à l'immunité, l'amélioration de la digestibilité de certains aliments, la synthèse de certains nutriments, la protection contre certains pathogènes, la communication intra- et inter-spécifique (Chatterjee *et al.*, 2017). Le tract intestinal de plusieurs espèces d'insectes est réputé être composé de communautés microbiennes très diversifiées (Engel & Moran, 2013; Mrazek *et al.*, 2008; Popa *et al.*, 2012), ces dernières jouant d'importants rôles dans les mécanismes liés aux différentes voies métaboliques de l'organisme hôte, à l'utilisation de différents polymères, à la méthanogenèse, à la dégradation de pesticides, ou encore à la protection contre certains pathogènes (Chatterjee *et al.*, 2017; Fox-Dobbs *et al.*, 2010; Nardi *et al.*, 2002). Le **tableau 0.2** récapitule quelques fonctions documentées de microorganismes associés aux insectes. Les insectes sont considérés comme des vecteurs d'un microbiote pouvant leur permettre de survivre et de s'adapter à leur environnement et dont la structure et les fonctions peuvent être modulées par des pressions stochastiques et déterministes exercées par des facteurs biotiques et abiotiques. Cette association essentielle porte le nom d'association symbiotique. Le concept de symbiose correspondant à une association entre des individus de différentes espèces sur une portion significative de leurs vies, les participants à cette association étant des symbiontes (Brownlie & Johnson, 2009; Margulis, 1992). Cette association qui peut être mutualiste (bénéfique), commensale (neutre) ou même parasitaire (nuisible) suggère des interactions potentielles entre les protagonistes impliqués dans l'association. Si de nombreuses études

rapportent les effets bénéfiques du microbiote pour leurs hôtes, des travaux de recherche ont également pointé la stabilité et la sensibilité de ce dernier lorsqu'il est confronté à des facteurs environnementaux pouvant aller jusqu'à entraver les fonctions initiales de l'hôte. Au regard de l'importance du microbiote pour son hôte et vis-versa, certaines études suggèrent que le microbiote et son hôte pourraient être considérés comme une unité écologique à part entière appelée « holobionte » (Gilbert *et al.*, 2012; Zilber-Rosenberg & Rosenberg, 2008). Le **tableau 0.2** présente quelques fonctions attribuées aux microorganismes associés aux insectes.

Tableau 0.2 Tableau récapitulatif de quelques fonctions attribuées aux microorganismes associés aux insectes.

Espèces	Hôte	Rôles	Références
Candida spp.	*Phoracantha semipunctata*	Dégradation de plusieurs des sucres Implication dans la production de vitamines Dégradation de cellulose, pectine et glucoside	Chararas and Chipoulet (1983)
Cellulomonas xylanilytica	*Dendroctonus valens*	Dégradation de la cellulose	Morales-Jiménez *et al.* (2009)
Champignons ophiostomatoïdes	*Dendroctonus* spp. *D. ponderosae* & *D. rufipennis* *D. ponderosae*	Implication dans la modulation du système de défense de l'arbre Implication potentielle dans la production de stérols pour les insectes Implication dans la production de composés organiques volatils (VOC)	Lieutier *et al.* (2009) Bentz and Six (2006) Cale *et al.* (2016)
Ogataea pini	*Dendroctonus brevicomis*	Production de composés volatils inhibant *Beauveria bassianna* Contribution à l'augmentation la croissance de (les composés volatils de l'espèce *O. pini* auraient des effets positifs sur les champignons mutualistes et un effet antagoniste sur les champignons pathogènes) Contribution à la dégradation de certains terpènes	Davis *et al.* (2011)
Pseudomonas spp.	*D. armandi*	Activité cellulolytique	Hu *et al.* (2014)
Pseudomonas fluorescens	*D. rhizophagus*	Implication à la dégradation des terpènes et des composés phénoliques	Morales-Jiménez *et al.* (2012)
Pseudomonas viridiflava	*D. rhizophagus*	Activité antifongique contre des champignons antagonistes	Cardoza *et al.* (2006)
Pseudomonas, Serratia et *Rahnella aquatilis*	*D. valens*	Contribution à la conversion de composés volatils (verbanol → verbanone)	Xu *et al.* (2015)
Pseudomonas et *Rahnella*	*D. ponderosae*	Implication dans la dégradation des monoterpènes	Adams *et al.* (2013)
Rahnella aquatilis	*D. valens* *D. rhizophagus*	Inhibition de la croissance de champignons ophiostomatoides : *O. montium* et *G. clavigera* Fixation et concentration de l'azote	Morales-Jiménez *et al.* (2009) Morales-Jiménez *et al.* (2012)
Serratia spp.	*D. ponderosae*	Implication dans l'inhibition de croissance des champignons ophiostomatoides (*O. montium* et *G. clavigera*)	Winder *et al.* (2010)
Serratia grimesii	*D. rufipennis*	Activité antifongique contre des champignons antagonistes	Cardoza *et al.* (2006)
Serratia marcescens	*D. ponderosae*	Implication dans la dégradation des terpènes	Boone *et al.* (2013)

1.3.1 Facteurs pouvant causer la variabilité du microbiome

Les relations symbiotiques entre les insectes et leur microbiote correspondent à des interactions constamment modulées par différents facteurs, dont la phytochimie de l'arbre et la présence d'un microbiote propre aux feuilles de l'arbre. Ainsi, les associations des microorganismes associés aux insectes, notamment à leur tract intestinal, peuvent être prédites par des mécanismes d'assemblage stochastique et/ou déterminisme (Brooks *et al.*, 2016). Parmi les facteurs pouvant influencer la structure du microbiome des insectes, on peut distinguer les facteurs biotiques (tels que la composition de la communauté microbienne et surtout aux interactions qui se font entre

les différents membres de cette communauté) des facteurs abiotiques (tels que la température, l'altitude, la position géographique, l'organisme hôte, l'alimentation). La température représente aussi un facteur très important pouvant impacter sur les communautés microbiennes parce qu'elle influence plusieurs mécanismes tels que la rigidité (diminution de la fluidité) et la fluidité membranaire, la diminution de l'activité des enzymes (diminution de la croissance cellulaire), augmentation de l'activité des métabolites (augmentation de la croissance cellulaire), la dénaturation des protéines (arrêt de la croissance cellulaire), augmentation de la perméabilité membranaire (transport et synthèse de l'ATP) et la modulation du pH (effets sur la transcription de l'ADN, la synthèse des protéines et l'activité enzymatique) (Jordan & Tomberlin, 2017). De la même façon, la disponibilité de l'eau est un facteur essentiel pour l'accomplissement de la plupart des voies métaboliques des microorganismes et leur survie. Une étude réalisée avec un échantillonnage effectué sur 50 sites distribués selon un gradient latitudinal entre la Colombie-Britannique et l'ouest de l'Alberta a suggéré que les niveaux d'élévation des paysages pourraient impacter sur les communautés symbiotiques chez le *Dendroctonus ponderosae* Hopkins (Roe *et al.*, 2011). Le régime alimentaire semble avoir un impact significatif sur les communautés bactériennes associées au tract intestinal de plusieurs espèces d'insectes étudiées à ce jour : *Anthonomus grandis* Boheman (Ben Guerrero *et al.*, 2016), *Dryophthoridae* (Montagna *et al.*, 2015a), *Cerambycidae spp.* (Kim *et al.*, 2017).

D'une part, il a été rapporté que les plantes hôtes possèdent des mécanismes de défense générant des composés potentiellement néfastes pour les insectes agresseurs. Les travaux de Chen and Poland (2009) ont effectivement permis de mettre en évidence l'impact de l'âge de la feuille, de l'exposition à la lumière et de l'annélation sur la qualité nutritionnelle des feuilles et sur les composés de défenses de l'arbre, suggérant un effet significatif de ces trois paramètres sur les performances des adultes de l'agrile du frêne et leur préférence alimentaire. Cette préférence serait dictée par les concentrations totales en protéines et en hydrates de carbone non structuraux totaux (TNC) dans les feuilles (Chen & Poland, 2009). Ces mêmes travaux suggèrent également de bonnes capacités pour l'excrétion ou pour la détoxification phénolique pour les agriles adultes. D'autre part, d'autres travaux ont suggéré que le microbiote de la phyllosphère interagirait avec celui de l'insecte au moment de l'attaque de l'arbre (Wielkopolan & Obrepalska-Steplowska, 2016). Les travaux de Durand *et al.* (2015) suggéraient que la relation tripartite et symbiotique qu'entretiennent les insectes avec leurs champignons et bactéries associés pourrait être modulée par des facteurs environnementaux.

1.4 Hypothèse générale – objectifs – Approche expérimentale

1.4.1 Hypothèse générale

Malgré l'existence de stratégies de lutte actuelle pour tenter de maitriser la propagation de l'agrile du frêne en Amérique du Nord, il reste très difficile de détecter des infestations initiales dans un délai raisonnable. Si les insectes entretiennent avec leur microbiote une relation intime et essentielle à la survie de l'insecte, il est légitime de penser que le microbiote de l'agrile du frêne pourrait être sensible aux changements environnementaux que l'insecte doit affronter au moment de son établissement. Les travaux de Bergeron (2016) portant sur la caractérisation taxonomique des communautés bactériennes associées à l'agrile du frêne avaient suggéré une variabilité du microbiome de l'agrile qui aurait une relation avec le niveau d'établissement de l'insecte. Il existe très peu de connaissances disponibles à ce jour sur la contribution potentielle du microbiome intestinal des insectes ravageurs dans leur mécanisme d'adaptation dans son nouvel environnement. Ainsi, l'hypothèse de travail de la présente thèse est que la structure taxonomique du microbiome du tract intestinal de l'agrile du frêne (communautés bactérienne et fongique) serait modulée par des facteurs environnementaux, notamment certaines propriétés des arbres hôtes.

1.4.2 Objectifs

L'objectif général de ce projet est de mettre en évidence les variations de la structure du microbiome du tract intestinal des adultes d'*A. planipennis* face à certaines propriétés de l'arbre hôte, notamment la densité de la population d'insectes présente sur l'arbre hôte, les microorganismes associés à la phyllosphère ou encore la phytochimie des arbres hôtes soumis à l'attaque par les herbivores. Ainsi, trois objectifs spécifiques ont été fixés : **(1)** évaluer l'effet de la taille de la population d'insectes adultes dans l'arbre hôte sur le microbiome de l'insecte ; **(2)** évaluer l'effet des propriétés intrinsèques de l'arbre hôte sur la structure du microbiome de l'insecte ; et **(3)** explorer la complexité des relations microbes-microbes dans le tract intestinal de l'insecte. L'originalité de ces travaux réside dans l'approche globale mettant en évidence la nécessité d'inclure dans l'interprétation des résultats sur le microbiome et l'existence de plusieurs autres paramètres pouvant exercer une influence discrète, mais non négligeable sur les processus mécaniques de la plasticité du microbiome. Elle permet de donner un peu plus de profondeur dans l'interprétation et surtout d'avoir une meilleure compréhension de la complexité des processus mécaniques du microbiome des insectes.

1.4.3 Approche expérimentale

L'ensemble des travaux de cette thèse visait à établir un portrait global des communautés microbiennes associées au tract intestinal de l'agrile du frêne adulte en utilisant une approche moléculaire basée sur l'analyse de l'ADN des microorganismes présents dans un échantillon. Ainsi, l'approche taxonomique qui permet de dresser un inventaire des taxons ou espèces présentes dans le milieu avec leurs abondances relatives a été utilisée. L'approche taxonomique s'appuie sur le gène codant pour l'ARN ribosomal (ARNr) en ciblant des régions dont la variation inter et intraspécifique est la plus conservée et qui permettent la différenciation des espèces. Ainsi, la région de l'espaceur interne transcrit (ITS) a été ciblée pour les communautés fongiques et la région v6-v8 pour les communautés bactériennes (**Figure 0.3**). Cette approche utilise le séquençage à haut débit d'amplicons PCR avec la technologie Illumina MiSeq. L'utilisation de l'outil bio-informatique a permis de filtrer le contenu, de caractériser et de déterminer les microorganismes présents dans les échantillons ainsi que leurs abondances. Les étapes de filtrage et de contrôle qualité sont importantes, car les techniques de séquençage peuvent générer des erreurs de séquençage (erreurs d'accumulation sur certaines fins de séquences), la duplication de certaines séquences, ou encore un manque d'uniformité de la couverture des séquences (Caporaso *et al.*, 2010; Edgar, 2010a; Knief, 2014; Schloss *et al.*, 2009). Ainsi, le traitement bio-informatique a été réalisé en s'appuyant sur l'approche proposée par (Edgar, 2010b; Edgar, 2016; Edgar & Flyvbjerg, 2015) dans le logiciel USEARCH, v10.0.240 (Edgar, 2010a). Finalement, de nombreuses approches statistiques ont été développées et adaptées à l'analyse des données génomiques, permettant ainsi d'apporter des éléments de réponses aux questions de recherche en écologie microbienne (Borcard *et al.*, 2018; Borcard *et al.*, 1992; De Cáceres *et al.*, 2010; Kembel *et al.*, 2011; Legendre & Legendre, 2012; Peres-Neto *et al.*, 2006). Ainsi, une analyse ciblée des communautés microbiennes obtenues a permis grâce à des choix adaptés d'outils statistiques, d'établir des faits et résultats liés à la problématique d'écologie microbienne que constitue notre hypothèse.

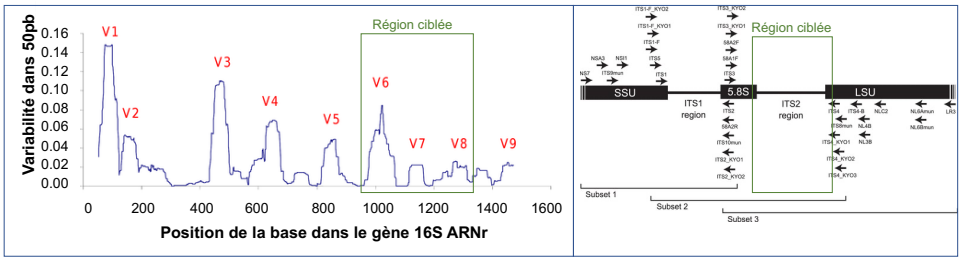


Figure 0.3 **Les régions ciblées par le séquençage, pour la communauté bactérienne, à gauche adaptation de Bodilis *et al.* (2012) et pour la communauté fongique, à droite adaptation de Toju *et al.* (2012)**

Les échantillonnages ont été effectués sur de nombreux sites variés s'étendant sur Montréal et sa grande région. Un premier échantillonnage d'insectes adultes dans les arbres hôtes (frênes) a révélé un gradient de niveaux d'infestations, permettant ainsi d'évaluer l'effet de la taille de la population d'insectes dans l'arbre hôte sur la structure taxonomique du microbiome du tract intestinal de l'agrile (communautés bactérienne et fongique). Un deuxième échantillonnage de plus grande envergure a permis d'évaluer la variation de ce microbiome sous la pression des propriétés intrinsèques à l'arbre hôte (microbiome de la phyllosphère, phytochimie des feuilles, et positionnement géographique). Après échantillonnage, chaque insecte choisi de manière aléatoire était traité selon les étapes suivantes : suppression des ailes et élytres ; lavage dans de l'éthanol ; rinçage dans de l'eau d'osmose ; dissection afin de récupérer le tract intestinal ; extraction de l'ADN ; séquençage ; traitement des données de séquençage ; analyses multivariées (Figure 0.4).

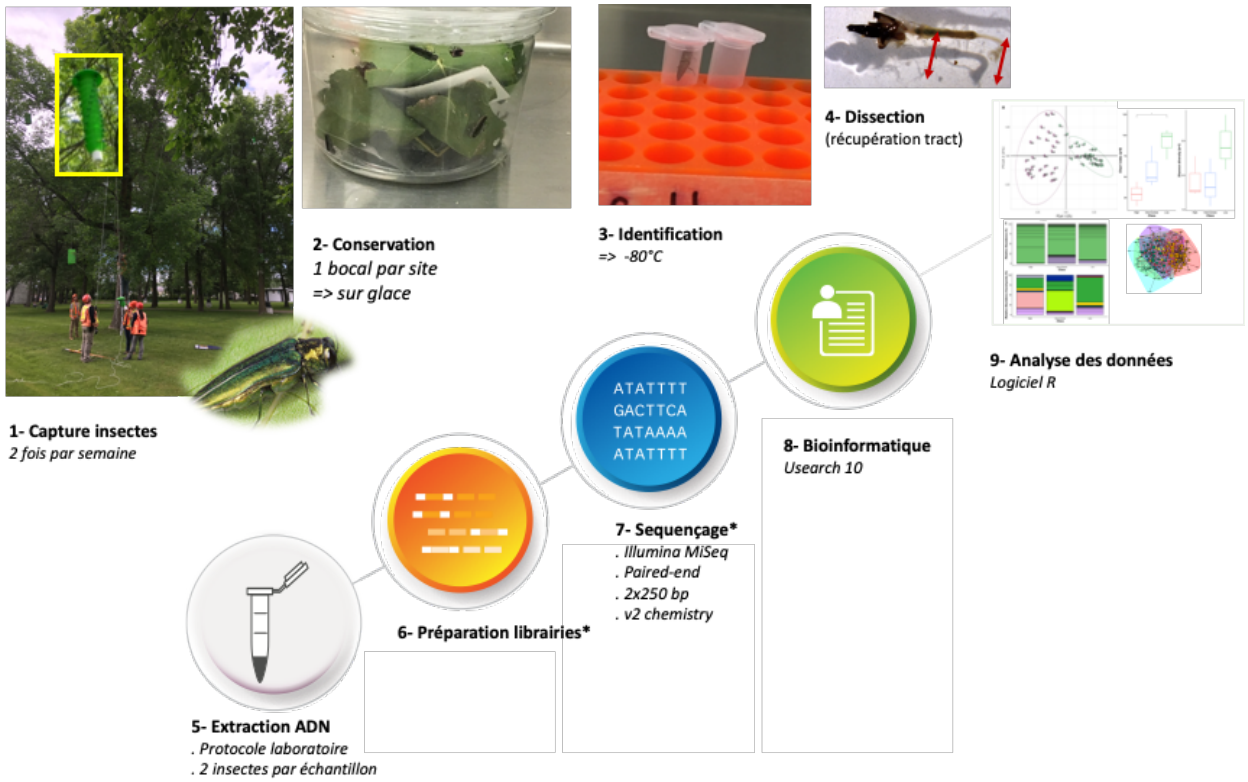


Figure 0.4 **Approche expérimentale globale, * étapes réalisées au Centre d'innovation Génome Québec.**

BIBLIOGRAPHIE (INTRODUCTION)

Adams AS, Aylward FO, Adams SM, Erbilgin N, Aukema BH, Currie CR, Suen G, Raffa KF (2013) Mountain pine beetles colonizing historical and naive host trees are associated with a bacterial community highly enriched in genes contributing to terpene metabolism. *Applied and environmental microbiology* 79(11):3468-3475.

Anulewicz AC (2006) Field surveys of emerald ash borer (Agrilus planipennis Fairmaire)(Coleoptera: Buprestidae) density and canopy dieback in three North American ash species. *The 2006 ESA Annual Meeting, December 10-13, 2006.*

Arimura G-i, Matsui K, Takabayashi J (2009) Chemical and molecular ecology of herbivore-induced plant volatiles: proximate factors and their ultimate functions. *Plant and Cell Physiology* 50(5):911-923.

Aukema JE, Leung B, Kovacs K, Chivers C, Britton KO, Englin J, Frankel SJ, Haight RG, Holmes TP, Liebhold AM (2011) Economic impacts of non-native forest insects in the continental United States. *PLoS One* 6(9):e24587.

Baranchikov Y, Mozolevskaya E, Yurchenko G, Kenis M (2008) Occurrence of the emerald ash borer, Agrilus planipennis in Russia and its potential impact on European forestry. *EPPO bulletin* 38(2):233-238.

Bellamy CL (2002) *Coleoptera: Buprestoidea.* CSIRO PUBLISHING,

Ben Guerrero E, Soria M, Salvador R, Ceja-Navarro JA, Campos E, Brodie EL, Talia P (2016) Effect of Different Lignocellulosic Diets on Bacterial Microbiota and Hydrolytic Enzyme Activities in the Gut of the Cotton Boll Weevil (Anthonomus grandis). *Frontiers in Microbiology* 7:2093.

Bentz BJ & Six DL (2006) Ergosterol content of fungi associated with Dendroctonus ponderosae and Dendroctonus rufipennis (Coleoptera: Curculionidae, Scolytinae). *Annals of the Entomological Society of America* 99(2):189-194.

Berg G, Rybakova D, Fischer D, Cernava T, Vergès M-CC, Charles T, Chen X, Cocolin L, Eversole K, Corral GH (2020) Microbiome definition re-visited: old concepts and new challenges. *Microbiome* 8(1):1-22.

Bergeron A (2016) *Caractérisation de la communauté bactérienne associée à l'agrile du frêne, Agrilus planipennis Fairmaire (Coleoptera: Buprestidae)*. (Université du Québec, Institut National de la Recherche Scientifique).

Bodilis J, Nsigue-Meilo S, Besaury L, Quillet L (2012) Variable copy number, intra-genomic heterogeneities and lateral transfers of the 16S rRNA gene in Pseudomonas. *PloS one* 7(4):e35647.

Boone CK, Keefover-Ring K, Mapes AC, Adams AS, Bohlmann J, Raffa KF (2013) Bacteria associated with a tree-killing insect reduce concentrations of plant defense compounds. *Journal of Chemical Ecology* 39(7):1003-1006.

Borcard D, Gillet F, Legendre P (2018) *Numerical ecology with R.* Springer,

Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology* 73(3):1045-1055.

Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR (2016) Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. *PLOS Biology* 14(11):e2000225.

Brownlie JC & Johnson KN (2009) Symbiont-mediated protection in insect hosts. *Trends in microbiology* 17(8):348-354.

Burr SJ & McCullough DG (2014) Condition of green ash (Fraxinus pennsylvanica) overstory and regeneration at three stages of the emerald ash borer invasion wave. *Canadian Journal of Forest Research* 44(7):768-776.

Cale JA, Taft S, Najar A, Klutsch JG, Hughes CC, Sweeney JD, Erbilgin N (2016) Mountain pine beetle (Dendroctonus ponderosae) can produce its aggregation pheromone and complete brood development in naïve red pine (Pinus resinosa) under laboratory conditions. *Canadian Journal of Forest Research* 45(12):1873-1877.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 7(5):335-336.

Cappaert D, McCullough DG, Poland TM, Siegert NW (2005b) Emerald ash borer in North America: a research and regulatory challenge.

Cardoza YJ, Klepzig KD, Raffa KF (2006) Bacteria in oral secretions of an endophytic insect inhibit antagonistic fungi. *Ecological Entomology* 31(6):636-645.

CFIA (2020) Emerald Ash Borer - Agrilus planipennis. *Canadian Food Inspection Agency Canada*.

Chakraborty S, Whitehill JGA, Hill AL, Opiyo SO, Cipollini DON, Herms DA, Bonello P (2014) Effects of water availability on emerald ash borer larval performance and phloem phenolics of Manchurian and black ash. *Plant, Cell & Environment* 37(4):1009-1021.

Chararas C & Chipoulet J-M (1983) Studies on the digestion of cellulose by the larvae of the Eucalyptus borer, Phoracantha semipunctata (Coleoptera: Cerambycidae). *Australian journal of biological sciences* 36(3):223-234.

Chatterjee S, Datta S, Sharma S, Tiwari S, Gupta DK (2017) Health and environmental applications of gut microbiome: a review. *Ecological Chemistry and Engineering S* 24(3):467-482.

Chen Y & Poland TM (2009) Interactive influence of leaf age, light intensity, and girdling on green ash foliar chemistry and emerald ash borer development. *J Chem Ecol* 35(7):806-815.

Chen Y, Whitehill JG, Bonello P, Poland TM (2011) Differential response in foliar chemistry of three ash species to emerald ash borer adult feeding. *Journal of chemical ecology* 37(1):29-39.

Davis TS, Hofstetter RW, Foster JT, Foote NE, Keim P (2011) Interactions between the yeast Ogataea pini and filamentous fungi associated with the western pine beetle. *Microbial Ecology* 61(3):626-634.

De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* 119(10):1674-1684.

Duan JJ, Bauer LS, Van Driesche RG (2017) Emerald ash borer biocontrol in ash saplings: The potential for early stage recovery of North American ash trees. *Forest Ecology and Management* 394:64-72.

Durand A-A, Bergeron A, Constant P, Buffet J-P, Déziel E, Guertin C (2015) Surveying the endomicrobiome and ectomicrobiome of bark beetles: The case of Dendroctonus simplex. 5:17190.

Edgar R (2010a) *Usearch.* Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), p

Edgar RC (2010b) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19):2460-2461.

Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv* :081257.

Edgar RC & Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 31(21):3476-3482.

Engel P & Moran NA (2013) The gut microbiota of insects‐diversity in structure and function. *FEMS microbiology reviews* 37(5):699-735.

Eyles A, Bonello P, Ganley R, Mohammed C (2010) Induced resistance to pests and pathogens in trees. *New Phytologist* 185(4):893-908.

Feldhaar H (2011) Bacterial symbionts as mediators of ecologically important traits of insect hosts. *Ecological Entomology* 36(5):533-543.

Fox-Dobbs K, Doak DF, Brody AK, Palmer TM (2010) Termites create spatial structure and govern ecosystem function by affecting N2 fixation in an East African savanna. *Ecology* 91(5):1296-1307.

Francese JA, Fraser I, Lance DR, Mastro VC (2011) Efficacy of multifunnel traps for capturing emerald ash borer (Coleoptera: Buprestidae): effect of color, glue, and other trap coatings. *Journal of economic entomology* 104(3):901-908.

Gandhi KJ & Herms DA (2010a) North American arthropods at risk due to widespread Fraxinus mortality caused by the alien emerald ash borer. *Biological Invasions* 12(6):1839-1846.

Gandhi KJ & Herms DA (2010b) Direct and indirect effects of alien insect herbivores on ecological processes and interactions in forests of eastern North America. *Biological Invasions* 12(2):389-405.

Gilbert SF, Sapp J, Tauber AI (2012) A symbiotic view of life: we have never been individuals. *The Quarterly review of biology* 87(4):325-341.

Haack RA, Jendek E, Liu H, Marchant KR, Petrice TR, Poland TM, Ye H (2002) The Emerald Ash Borer: A New Exotic Pest in North America.

Harborne A (1998) *Phytochemical methods a guide to modern techniques of plant analysis.* springer science & business media,

Harborne JB (1988) *Introduction to Ecological Biochemistry.* Academic Press, New-York, 3

Herms DA & McCullough DG (2014) Emerald ash borer invasion of North America: history, biology, ecology, impacts, and management. *Annu Rev Entomol* 59:13-30.

Hu X, Yu J, Wang C, Chen H (2014) Cellulolytic bacteria associated with the gut of Dendroctonus armandi larvae (Coleoptera: Curculionidae: Scolytinae). *Forests* 5(3):455-465.

Jendek E & Grebennikov V (2011) *Agrilus (Coleoptera, Buprestidae) of East Asia.* Nakladatelstvi Jan Farkač,

Jordan HR & Tomberlin JK (2017) Abiotic and biotic factors regulating inter-kingdom engagement between insects and microbe activity on vertebrate remains. *Insects* 8(2):54.

Kembel SW, Eisen JA, Pollard KS, Green JL (2011) The phylogenetic diversity of metagenomes. *PLoS One* 6(8):e23214.

Kim JM, Choi M-Y, Kim J-W, Lee SA, Ahn J-H, Song J, Kim S-H, Weon H-Y (2017) Effects of diet type, developmental stage, and gut compartment in the gut bacterial communities of two Cerambycidae species (Coleoptera). *Journal of Microbiology* 55(1):21-30.

Knief C (2014) Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Frontiers in plant science* 5.

Knight KS, Brown JP, Long RP (2013) Factors affecting the survival of ash (Fraxinus spp.) trees infested by emerald ash borer (Agrilus planipennis). *Biological Invasions* 15(2):371-383.

Kostova I & Iossifova T (2007) Chemical components of Fraxinus species. *Fitoterapia* 78(2):85-106.

Kovacs KF, Haight RG, McCullough DG, Mercader RJ, Siegert NW, Liebhold AM (2010) Cost of potential emerald ash borer damage in U.S. communities, 2009–2019. *Ecological Economics* 69(3):569-578.

Legendre P & Legendre LF (2012) *Numerical ecology.* Elsevier,

Lieutier F, Yart A, Salle A (2009) Stimulation of tree defenses by Ophiostomatoid fungi can explain attack success of bark beetles on conifers. *Annals of Forest Science* 66(8):801-801.

Liu H, Bauer LS, Gao R, Zhao T, Petrice TR, Haack RA (2004) Exploratory survey for the emerald ash borer, Agrilus planipennis (Coleoptera: Buprestidae), and its natural enemies in China. *Great Lakes Entomologist* 36(3-4)(Fall-Winter 2003).

Lyons DB, Caister C, De Groot P, Hamilton B, Marchant K, Scarr T, Turgeon J (2007) Survey guide for detection of emerald ash borer.

Lyons DB, Lavallée R, Kyei‐Poku G, Van Frankenhuyzen K, Johny S, Guertin C, Francese JA, Jones GC, Blais M (2012) Towards the Development of an Autocontamination Trap System to Manage Populations of Emerald Ash Borer (Coleoptera: Buprestidae) With the Native Entomopathogenic Fungus, Beauveria bassiana. *Journal of Economic Entomology* 105(6):1929-1939.

MacFarlane DW & Meyer SP (2005) Characteristics and distribution of potential ash tree hosts for emerald ash borer. *Forest Ecology and Management* 213(1):15-24.

Margulis L (1992) Biodiversity: molecular biological domains, symbiosis and kingdom origins. *Biosystems* 27(1):39-51.

Margulis L (1993) Origins of species: acquired genomes and individuality. *Biosystems* 31(2-3):121-125.

Marquis RJ (2004) Herbivores rule. *Science* 305(5684):619-621.

McCullough DG (2020) Challenges, tactics and integrated management of emerald ash borer in North America. *Forestry: An International Journal of Forest Research* 93(2):197-211.

McCullough DG, Siegert NW, Cappaert D, Poland TM, McDonald R (2007) Sinks, bark, and garlon: applied studies for emerald ash borer management. *Emerald ash borer and asian longhorhed beetle research and development review*. p 92.

Montagna M, Chouaia B, Mazza G, Prosdocimi EM, Crotti E, Mereghetti V, Vacchini V, Giorgi A, De Biase A, Longo S (2015a) Effects of the diet on the microbiota of the red palm weevil (Coleoptera: Dryophthoridae). *PLoS One* 10(1):e0117439.

Morales-Jiménez J, Zúñiga G, Ramírez-Saad HC, Hernández-Rodríguez C (2012) Gut-associated bacteria throughout the life cycle of the bark beetle Dendroctonus rhizophagus Thomas and Bright (Curculionidae: Scolytinae) and their cellulolytic activities. *Microbial Ecology* 64(1):268-278.

Morales-Jiménez J, Zúñiga G, Villa-Tanaca L, Hernández-Rodríguez C (2009) Bacterial community and nitrogen fixation in the red turpentine beetle, Dendroctonus valens LeConte (Coleoptera: Curculionidae: Scolytinae). *Microbial Ecology* 58(4):879-891.

Mrazek J, Strosova L, Fliegerova K, Kott T, Kopecny J (2008) Diversity of insect intestinal microflora. *Folia Microbiol (Praha)* 53(3):229-233.

Nardi JB, Mackie RI, Dawson JO (2002) Could microbial symbionts of arthropod guts contribute significantly to nitrogen fixation in terrestrial ecosystems? *J Insect Physiol* 48(8):751-763.

Orlova-Bienkowskaja MJ (2014) Ashes in Europe are in danger: the invasive range of Agrilus planipennis in European Russia is expanding. *Biological Invasions* 16(7):1345-1349.

Orlova-Bienkowskaja MJ, Drogvalenko AN, Zabaluev IA, Sazhnev AS, Peregudova EY, Mazurov SG, Komarov EV, Struchaev VV, Martynov VV, Nikulina TV (2020) Current range of Agrilus planipennis Fairmaire, an alien pest of ash trees, in European Russia and Ukraine. *Annals of Forest Science* 77(2):1-14.

Orlova-Bienkowskaja MJ & Bieńkowski AO (2016) The life cycle of the emerald ash borer Agrilus planipennis in European Russia and comparisons with its life cycles in Asia and North America. *Agricultural and Forest Entomology* 18(2):182-188.

Paré PW & Tumlinson JH (1999) Plant volatiles as a defense against insect herbivores. *Plant physiology* 121(2):325-332.

Peres-Neto PR, Legendre P, Dray S, Borcard D (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* 87(10):2614-2625.

Poland TM & McCullough DG (2006) Emerald ash borer: invasion of the urban forest and the threat to North America's ash resource. *Journal of Forestry* 104(3):118-124.

Popa V, Deziel E, Lavallee R, Bauce E, Guertin C (2012) The complex symbiotic relationships of bark beetles with microorganisms: a potential practical approach for biological control in forestry. *Pest Manag Sci* 68(7):963-975.

Pureswaran DS & Poland TM (2009) Host selection and feeding preference of Agrilus planipennis (Coleoptera: Buprestidae) on ash (Fraxinus spp.). *Environmental Entomology* 38(3):757-765.

Rodriguez-Saona C, Poland TM, Miller JR, Stelinski LL, Grant GG, Groot Pd, Buchan L, MacDonald L (2006) Behavioral and electrophysiological responses of the emerald ash borer, Agrilus planipennis, to induced volatiles of Manchurian ash, Fraxinus mandshurica. *CHEMOECOLOGY* 16(2):75-86.

Roe AD, Rice AV, Coltman DW, Cooke JEK, Sperling FAH (2011) Comparative phylogeography, genetic differentiation and contrasting reproductive modes in three fungal symbionts of a multipartite bark beetle symbiosis. *Molecular Ecology* 20(3):584-600.

Ryall KL, Fidgen JG, Turgeon JJ (2011) Detectability of the emerald ash borer (Coleoptera: Buprestidae) in asymptomatic urban trees by using branch samples. *Environmental Entomology* 40(3):679-688.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology* 75(23):7537-7541.

Schoonhoven LM, Van Loon B, van Loon JJ, Dicke M (2005) *Insect-plant biology.* Oxford University Press on Demand,

Siegert NW, McCullough DG, Liebhold AM, Telewski FW (2014) Dendrochronological reconstruction of the epicentre and early spread of emerald ash borer in N orth A merica. *Diversity and distributions* 20(7):847-858.

Slansky F & Haack R (1986) Age-specific flight behavior in relation to body weight and lipid content of Ips calligraphus reared in slash pine bolts with thick or thin inner bark (phloem). *Entomologia Experimentalis et Applicata* 40(2):197-207.

Srei N, Guertin C, Lavallée R, Lajoie M-È, Brousseau C, Bergevin R, Miller F, McMillin K, Trudel R (2020a) Microbial Control of the Emerald Ash Borer (Coleoptera: Buprestidae) Using Beauveria bassiana (Hypocreales: Cordycipitaceae) by the Means of an Autodissemination Device. *Journal of Economic Entomology* 113(6):2657-2665.

Srei N, Lavallée R, Guertin C (2020b) Horizontal transmission of the entomopathogenic fungal isolate INRS-242 of Beauveria bassiana (Hypocreales: Cordycipitaceae) in Emerald Ash Borer, Agrilus planipennis (Coleoptera: Buprestidae). *Journal of Economic Entomology* 113(1):543-545.

Toju H, Tanabe AS, Yamamoto S, Sato H (2012) High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PloS one* 7(7):e40863.

USDA (2009) Emerald Ash Borer. *Michigan State University Michigan USA*.

Valenta V, Moser D, Kapeller S, Essl F (2017) A new forest pest in Europe: a review of Emerald ash borer (Agrilus planipennis) invasion. *Journal of Applied Entomology* 141(7):507-526.

Villari C, Herms DA, Whitehill JG, Cipollini D, Bonello P (2016) Progress and gaps in understanding mechanisms of ash tree resistance to emerald ash borer, a model for wood-boring insects that kill angiosperms. *New Phytologist* 209(1):63-79.

Villari C, Whitehill JG, Cipollini D, Herms DA, Bonello P (2014) Mechanisms of Ash Resistance to Emerald Ash Borer: Progress and Gaps.

Wang, Yang Z, Gould J, Zhang Y, Liu G, Liu E (2010) The biology and ecology of the emerald ash borer, Agrilus planipennis, in China. *J Insect Sci* 10:128.

Wei X, Reardon D, Wu Y, Sun J-H (2004) Emerald ash borer, Agrilus planipennis Fairmaire (Coleoptera: Buprestidae), in China: a review and distribution survey. *Acta Entomologica Sinica* 47(5):679-685.

Wei X, Wu Y, Reardon R, SUN TH, Lu M, SUN JH (2007) Biology and damage traits of emerald ash borer (*Agrilus planipennis* Fairmaire) in China. *Insect Science* 14(5):367-373.

Wielkopolan B & Obrepalska-Steplowska A (2016) Three-way interaction among plants, bacteria, and coleopteran insects. *Planta* 244(2):313-332.

Winder RS, Macey DE, Cortese J (2010) Dominant bacteria associated with broods of mountain pine beetle, Dendroctonus ponderosae (Coleoptera: Curculionidae, Scolytinae). *Journal of the Entomological Society of British Columbia* 107:43-56.

Wu J (2016) *Transcriptomics and Gene Regulation.* Springer,

Wylie FR & Speight MR (2012) *Insect pests in tropical forestry.* CABI, Oxfordshire, UK

Xu L, Lou Q, Cheng C, Lu M, Sun J (2015) Gut-associated bacteria of Dendroctonus valens and their involvement in verbenone production. *Microbial ecology* 70(4):1012-1023.

Zilber-Rosenberg I & Rosenberg E (2008) Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS microbiology reviews* 32(5):723-735.

1 GUT MICROBIOME OF THE EMERALD ASH BORER, *AGRILUS PLANIPENNIS* FAIRMAIRE, AND ITS RELATIONSHIP WITH INSECT POPULATION DENSITY

Titre français : Le microbiome de l'agrile du frêne (*Agrilus Planipennis* Fairmaire) et sa relation avec la densité de population d'insectes

Auteurs :

Judith Mogouong[1] , Philippe Constant[1] , Robert Lavallée[2] , and Claude Guertin[1]✉

Affiliations professionnelles

[1]Institut national de la recherche scientifique, Centre Armand-Frappier Santé Biotechnologie, Laval, QC, Canada, H7V 1B7

[2]Natural Resources Canada, Canadian Forest Service, Quebec, QC, Canada, G1V 4C7

✉ Auteur correspondant

Titre de la revue ou de l'ouvrage :

FEMS Microbiology Ecology

Soumis le 6 février 2020, accepté le 9 juillet 2020, publié le 13 juillet 2020

doi : 10.1093/femsec/fiaa141

Contribution des auteurs :

Judith Mogouong : Conception et planification du design expérimental, échantillonnage sur le terrain, réalisation des expériences et des analyses statistiques, analyse des résultats, préparation des figures, rédaction du manuscrit, révision du manuscrit

Philippe Constant : Conception et planification du design expérimental, réalisation des expériences, révision du manuscrit

Robert Lavallée : Conception et planification du design expérimental, échantillonnage sur le terrain, révision du manuscrit

Claude Guertin : Conception et planification du design expérimental, réalisation des expériences, préparation des figures, analyse des résultats, révision du manuscrit

1.1 Résumé en français

Les communautés microbiennes associées au tract intestinal de coléoptères jouent un rôle crucial dans leurs capacités d'adaptation. Des facteurs environnementaux tels que la température ou la nutrition affectent naturellement le microbiome de l'insecte, mais un changement des conditions locales comme la densité de population sur un arbre hôte pourrait également conduire à des changements dans le microbiote. L'agrile du frêne (AGR), *Agrilus planipennis* Fairmaire, est un foreur du bois exotique qui cause des dommages environnementaux et économiques aux frênes en Amérique du Nord. Cette étude visait à décrire la structure taxonomique du microbiome intestinal de l'agrile du frêne et à explorer sa relation potentielle avec la taille de la population d'insectes. Le nombre d'adultes de l'agrile du frêne recueillis par arbre sur un transect de 75 km à partir d'un épicentre a permis la création de classes distinctes de densité de population. Les Gammaprotéobactéries et Ascomycota prédominent respectivement les communautés bactériennes et fongiques, comme déterminé par séquençage du gène de l'ARNr 16S bactérien et de l'espaceur transcrit interne fongique ITS2. La richesse en espèces et la diversité de la communauté bactérienne ont montré une dépendance significative de la densité de population. De plus, l'analyse de la diversité α et de la diversité β a révélé certains variants de séquence d'amplicons indicateurs suggérant que la plasticité du microbiome intestinal pourrait être liée à la densité des populations de l'agrile du frêne chez les arbres hôtes.

1.2 Abstract

The gut microbial communities of beetles play crucial roles in their adaptive capacities. Environmental factors such as temperature or nutrition naturally affect the insect microbiome, but a shift in local conditions like the population density on a host tree could also lead to changes in the microbiota. The emerald ash borer (EAB), Agrilus planipennis Fairmaire, is an exotic wood borer that causes environmental and economic damage to ash trees in North America. This study aimed to describe the taxonomic structure of the EAB gut microbiome and explore its potential relationship with borer population size. The number of EAB adults collected per tree through a 75 km transect from an epicenter allowed the creation of distinct classes of population density. The Gammaproteobacteria and Ascomycota predominated in bacterial and fungal communities respectively, as determined by sequencing of the bacterial 16S rRNA gene and the fungal internal transcribed spacer ITS2. Species richness and diversity of the bacterial community showed significant dependence on population density. Moreover, α-diversity and β-diversity analysis revealed some indicator amplicon sequence variants suggesting that the plasticity of the gut microbiome could be related to the EAB population density in host trees.

1.3 Introduction

Microorganisms that colonize insects play an essential role in the health of their hosts by helping them with niche occupation and competition, and in many functional processes, such as pathogen defense, digestion, and nutrient absorption (Douglas, 2015; Engel & Moran, 2013). Microbial symbiosis has been extensively studied and documented (Brooks *et al.*, 2016; Hurst, 2017; Paracer & Ahmadjian, 1999; Popa *et al.*, 2012). Insect guts are environments for specific microbial colonization, and it has been shown that some gut microorganisms contribute to food digestion through metabolism of indigestible plant polymers (Engel & Moran, 2013) or by producing essential vitamins and other nutrients for their hosts (Douglas, 2009). Many studies have shown interactions between symbionts and hosts (Nikoh *et al.*, 2011; Panteleev *et al.*, 2007; Zindel *et al.*, 2011). For example, xylophagous insects require symbiotic microorganisms with high biosynthetic abilities to compensate for the low levels of amino acids in their diet, whereas leaf-eaters need bacteria and fungi with ability to break down molecules such as cellulose (Hurst, 2017).

The emerald ash borer (EAB), *Agrilus planipennis* Fairmaire, is an important invasive pest detected for the first time in North America in 2002 (Cappaert *et al.*, 2005b; Haack *et al.*, 2002; Wang *et al.*, 2010). Native to Asia, the EAB usually attacks stressed, weakened, or dying trees (Wei *et al.*, 2007), but in some cases it may also harm healthy ash trees (Wang *et al.*, 2010). The EAB has caused the greatest economic and biological impacts of any exotic insect pest (Aukema *et al.*, 2011; Cappaert *et al.*, 2005b; Herms & McCullough, 2014; Kovacs *et al.*, 2010) and represents a severe threat to the ash tree population in North America (Gandhi & Herms, 2010a; Herms *et al.*, 2004; Wei *et al.*, 2007). Despite strategies implemented by the United States and Canadian governments to limit the spread of this pest (Duan *et al.*, 2012; Herms & McCullough, 2014; Poland *et al.*, 2005), the EAB is now found in thirty-five American states and five Canadian provinces (Emerald ash Borer information network, 2020). Infestations of newly attacked ashes are challenging to detect, and symptoms generally do not appear until the trees are severely infested (Lyons *et al.*, 2007). Among the strategies used to control the pests' spread, methods based on microbiological approaches are expected to increase because they are safer than using chemical insecticides (Chattopadhyay *et al.*, 2017; Mascarin & Jaronski, 2016; Starnes *et al.*, 1993). Although the number of studies on the relationship between a pest's microbiome and its environment has increased, the information related to the EAB microbiome remains limited. The main goal of this study was to generate some relevant knowledge that could be helpful in biological pest control.

A previous study, both culture-dependent and culture-independent, has revealed some variations in gut microorganisms associated with the development and life cycle of EAB (Vasanthakumar *et al.*, 2008). Some other results suggested that bacterial and fungal profiles could be related significantly to the developmental stage of the insect hosts (Delalibera *et al.*, 2007; Huang & Zhang, 2013; Morales-

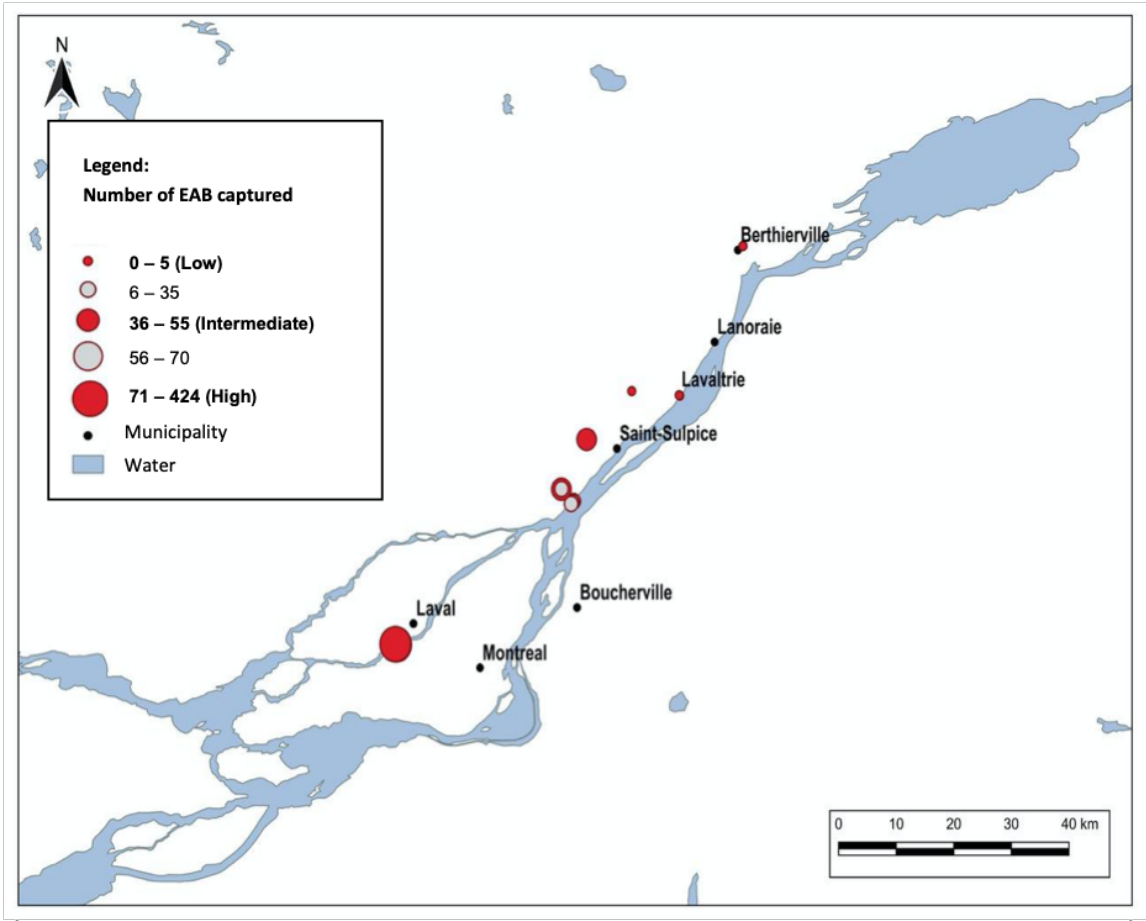
Jiménez *et al.*, 2012; Rani *et al.*, 2009), to the nutritional status (Broderick, 2004; Douglas, 2009; Medina *et al.*, 2011; Montagna *et al.*, 2015b; Priya *et al.*, 2012), and even to the altitude of their habitat (Montagna *et al.*, 2015a; Roe *et al.*, 2011). Zhao *et al.* (2013) suggested that it might be possible to predict the distribution and population density of a nematode, *Bursaphelenchus xylophilus* (Steiner & Buhrer), in newly introduced areas by using population models of multispecies interactions related to the symbiosis between that nematode and the population of the Japanese sawyer beetle, *Monochamus alternatus* Hope. Herbivory is known to induce response cascades in plants such as the production of elicitors (Bonaventure *et al.*, 2011) and changes in secondary metabolites (Rosenthal & Berenbaum, 2012) and hormones (Howe & Schaller, 2008). The gut microbiota of adult EABs could also be indirectly affected by the size of the insects' population. Steinhaus (1958) suggested that insect crowding could promote bacterial or viral diseases. To the best of our knowledge there have been no published studies of the potential relationship between an insect's population size and its gut microbiome. We hypothesize that the variation in the number of EAB adults collected per host ash tree could have an impact on the taxonomical structure of its gut-associated microorganisms. An amplicon sequencing approach was used to (a) determine the composition of gut bacterial and fungal communities of EABs, (b) to analyze the variation of taxonomical structures of microbes based on the insects' population density per tree, and (c) to identify microbial taxa associated with insect density classes as indicator amplicon sequence variants (ASVs).

1.4 Materials and Methods

1.4.1 Site location and insect collection and sampling

Twenty-one ash trees were randomly selected across a 75 km northeast transect along the St. Lawrence River from the Montreal (Qc, Canada) metropolitan area, which was identified as the northeast limit of EAB infestation in 2016 (**Figure 1.1**). EAB adults were collected using green Lindgren 12-unit multiple-funnel traps (Synergy Semiochemicals, Burnaby, BC, Canada) installed at one third of the apical part of selected trees using a simple halyard system as described by Hughes *et al.* (2014) to facilitate monitoring. For each trap, leaflets of the host tree were put into the dry container as substrate and food source for the beetles. Based on personal observations on adult flight activity, and the calculation method of degree days described by Snyder (1985), traps were deployed and monitored when a total of 310 degree-days (base 10) was reached. EAB adults were collected every two to three days between June 20th and August 10th, 2016. When collected, each container was identified and put on ice during its transport to the laboratory. Adults were then individually transferred into identified sterile 1.5 mL microfuge tubes and frozen at -80°C before DNA extraction. Using Sturges' rule (Scott, 2009), the EAB population densities were divided into five categories based on the distribution of the number of insects collected per tree. From these five population-density classes, three were selected for the subsampling

strategy to relate classes to the structure of the gut microbial communities (**Figure 1.1**). Three biological replicates represented each class. The three composite samples representative of low population density comprised seven trees, where less than five individuals per trap were collected (because of low numbers, insects collected from two to three trees were combined for each biological replicate). The intermediate density classes were associated with three ash trees on which 35 to 56 insects per trap were caught, and more than 70 insects per trap represented the high-density population (each biological replicate was assembled with insects collected from a single tree).



Classes	Samples	Sites	Latitude	Longitude	Number of EAB captured
Low	FAD1	Mic-07	45.7383837	-73.4415249	5
		Mic-17	45.8836197	-73.2754047	1
	FAD2	Mic-11	45.7347991	-73.4454175	5
		Mic-09	45.7378355	-73.4417817	2
	FAD3	Mic-16	45.8898493	-73.3509158	5
		Mic-18	45.8840697	-73.2763236	2
Intermediate	MOD1	Mic-12	45.7549191	-73.4606546	38
	MOD2	Mic-13	45.7549777	-73.4606284	39
	MOD3	Mic-15	45.8232803	-73.4212631	54
High	FOD1	Mic-03	45.5414861	-73.7193285	77
	FOD2	Mic-05	45.5407558	-73.7191667	105
	FOD3	Mic-06	45.5417561	-73.7201217	73

Figure 1.1 : Transect of the EAB sampling area along the St. Lawrence River.

The North-East transect from Montreal (QC) along the St. Lawrence River showing randomly selected ash trees located in a 75 km and a gradient of insect-densities captured using green Lindgren 12-unit multi-funnel traps. The number of captured insects per tree ranges from 1 to 424. Three distinct classes of insect-densities were used for statistical analyses. The circles indicate the sampling sites. The size of the circles increases with the number of insects collected per site. The circles filled in red are the sites considered in the study (some selected sites are hidden behind others because of the scaling), whereas those filled in grey represent the sites not considered in the study.

1.4.2 **DNA extraction**

Preliminary results using our DNA extraction protocol showed that the gut contents from five adult EABs were required to recover enough microbial DNA for downstream PCR amplification. Therefore, each replicate was formed of five randomly selected insects from individual trees belonging to the same population density class, except for the Low-density class for which we used a combination of insects from trees belonging to the same class because of the insufficient number of insects collected per tree. For each class, DNA extraction was performed on three biological replicates. The beetle's guts were recovered by first removing the wings of each insect using sterile tweezers and scissors. The exoskeleton was then sterilized by vortexing (Fisher Vortex Genie 2, Ottawa, ON, Canada) the beetle in 1 mL of 70% ethanol for 1 min. After discarding the ethanol, the insects were individually rinsed with 1 mL of sterile water for 30 s using the vortexer. Dissection was performed in sterile phosphate-buffered saline (PBS). Tissue was opened up longitudinally and the whole gut was gently removed and placed in a sterile 1.5 mL tube containing 100 µL of extraction buffer (50 mM Tris-HCl, 5 mM EDTA-2Na, 3% SDS, pH 8.0). The guts were macerated using sterilized pestles, and total genomic DNA was extracted using the mechanical lysis method as described by Durand *et al.* (2015) A negative control containing all the extraction buffers but no insect parts was utilized, and no amplicons were observed. The DNA concentration was estimated using the Quant-iT^{TM} PicoGreen$^{®}$ dsDNA Assay Kit (Invitrogen, Life Technologies, Burlington, ON, Canada) following the manufacturer's instructions.

1.4.3 **PCR amplification, amplicon sequencing and data processing**

The V6-V8 region of the bacterial 16S rRNA gene was sequenced using the primers, B969F-CS1 (5'-ACGCGHNRAACCTTACC-3') and BA1406R-CS2 (5'- ACGGGCRGTGWGTRCAA-3') (Comeau *et al.*, 2011). The fungal ITS2 region was sequenced using the primers, ITS3_KYO2 (5'-GATGAAGAACGYAGYRAA-3') and ITS4_KYO3 (5'-CTBTTVCCKCTTCACTCG-3') (Toju *et al.*, 2012). For each population density class, three biological replicates were performed. The sequencing was done on an Illumina Miseq using 2 x 250 bp paired-ends at the Quebec Genome Innovation Centre. The raw sequences were processed by using a pipeline in the USEARCH10 software. The pipeline involved merging of paired-end reads (the total length of paired-ends was 400-500 bp for bacteria and 200-440 bp for fungi), quality filtering (the maximum overlap mismatch accepted was five bases or 10%), trimming of primers from merged reads and discarding of singletons and reads with >1 expected error (Edgar & Flyvbjerg, 2015), and dereplication to find unique sequences. The use of amplicon sequence variant (ASV) clustering with 100% sequence identity was chosen over OTU clustering (generally representing clusters with <100% sequence identity) because of its "reusability, its reproducibility, and its higher resolution" (Callahan *et al.*, 2017). After dereplication, the reads were denoised and the ASVs were clustered by considering reads with a minimum frequency of eight as representative sequences (Edgar,

2016). ASVs with frequencies <0.005% were discarded. The taxonomy assignment was completed using the Ribosomal Database Project (RDP, https://rdp.cme.msu.edu/classifier/) training set v.16 of 16S rRNA genes for bacteria and RDP Warcup training set v.2. for fungi (Cole *et al.*, 2014; Wang *et al.*, 2007). For the bacterial dataset, the reads assigned to archaea and chloroplasts were removed to focus only on the eubacterial community. At this step, the resulting ASV tables were used to assess biodiversity based on the Hill's numbers (q) estimators for species richness (q = 0), the exponential of Shannon index (q = 1), and the inverse of Simpson index (q = 2) using the package iNext version 2.0.19 (Hsieh *et al.*, 2016). ASV tables of bacteria and fungi were used for statistical analyses, including β-diversity and indicator species analysis. Raw reads are available in the Sequence Read Archive of the National Center for Biotechnology Information (SRA, NCBI) under the Bioproject number (PRJNA544945) and the SRA (SUB5673374).

1.4.4 **PCR detection of indicator species**

As some indicator ASVs were found specific to the low-density class, a pair of primers was designed to detect them. The representative ASV sequences were aligned using the ClustalW algorithm in the software Mega 5.2 (Kumar *et al.*, 2016) to identify consensus regions and select the following specific primers:

MOD1-F 5'- CGTCAGCTCGTGTTGTGAAA-3'

MOD1-R 5'- GCTTCTCTTTGTATGCGCC-3'.

The 50 µl PCR mixture contained 25 mM $MgCl_2$, 10 µg BSA, 10 mM dNTPs, 10 mM of each primer, 5 U Taq DNA polymerase and ThermoPol® buffer (New England Biolabs, Whitby, ON, Canada), 3 µL template DNA and nuclease-free water to obtain a final volume of 50 µL. The PCR assay consisted of an initial denaturation step at 94°C for 3 min, followed by 30 amplification cycles of denaturation at 94°C for 30 s, annealing at 59°C for 30 s and elongation at 72°C for 30 s), with a final extension at 72°C for 10 min. The assay was run on template DNA from the different insect density classes to test the specificity of the indicator ASVs.

1.4.5 **Statistical analysis**

All statistical analyses were performed using the R software (Team, 2013). One-way ANOVA followed by *post hoc* Tukey tests were then applied using the 'vegan' version 2.5-4 package (Oksanen *et al.*, 2007) to compare the α-diversity between the three EAB density classes. The core microbiome of EAB adults including taxa prevalent in 80% of samples was determined using the microbiome version 1.5.23 package (Lahti *et al.*, 2017). Hierarchical clustering of bacterial and fungal community profiles was established using the unweighted pair group method with arithmetic mean (UPGMA) in which significant

clusters were identified using the *simprof* function in the clustsig version 1.1 package (Whitaker & Christman, 2014). A principal coordinate analysis (PCoA) was performed on Unifrac distances (weighted and unweighted) using the *Cmdscale* function in the stats version 3.5.3 package (Venables *et al.*, 2002) to visualize the data distribution. To evaluate the proxy effect of insect density on gut microbiome patterns, a permutational multivariate analysis of variance based on distance matrices (PERMANOVA) was conducted using the Adonis2 function in the 'vegan' version 2.5-4 package (Oksanen *et al.*, 2007). Pairwise comparisons were performed using the Wilcoxon rank-sum test with Bonferroni correction to elucidate the effect of each insect density class on taxonomic patterns of the EAB microbiome based on PERMANOVA results. The same statistical analyses were performed using the Operational Taxonomic Unit (OTU) approach on clusters with 97% sequence identity. De Cáceres *et al.* (2010) defined indicator species as "ecological indicators of community or habitat types, environmental conditions, or environmental changes" and proposed their identification and use for microbiome studies. In our study the indicator species corresponded to the biological sequences (ASVs) associated with a specific group of samples, and we used the Indicspecies version 1.7.6 package (De Cáceres & Jansen, 2016) to identify the ASVs driving the differences in microbiome patterns amongst insect density classes with 9999 permutations. Based on the indicator ASVs found for the low population density class, a pair of primers was designed to target a consensus region of all the corresponding sequences, which allowed us to evaluate the discrimination between the three classes by independent PCRs under specific conditions. A Spearman correlation test was performed between the indicator abundance values and the number of insects collected per tree.

1.5 Results

The highest number of insects was caught in Montreal, decreasing along the St. Lawrence River transect. The variation of the number of captured insects along the transect was explained with the linear regression (**Figure Suppl.1.1**, Supporting Information): $N = 4.44–0.06 * d$.

Where N is the number of captured insects per tree and d is the distance from Montreal ($R^2 = 0.74$ and $p < 0.0001$).

1.5.1 Microbial communities comprising the gut microbiome of adult A. planipennis

After cluster analysis of 453 530 sequences and quality control, a total of 187 ASVs associated with the adult EAB gut bacterial community were obtained. The EAB gut fungi were represented by 514 813 sequences clustered into 71 ASVs (**Table Suppl. 1.1**, Supporting Information). A comparison of the observed rich- ness (Chao1 observed) with the richness estimated (Chao1 estimator) confirmed that the sequencing effort was sufficient to cover the diversity (81.5% coverage for bacteria and 90.7% for fungi; data not shown).

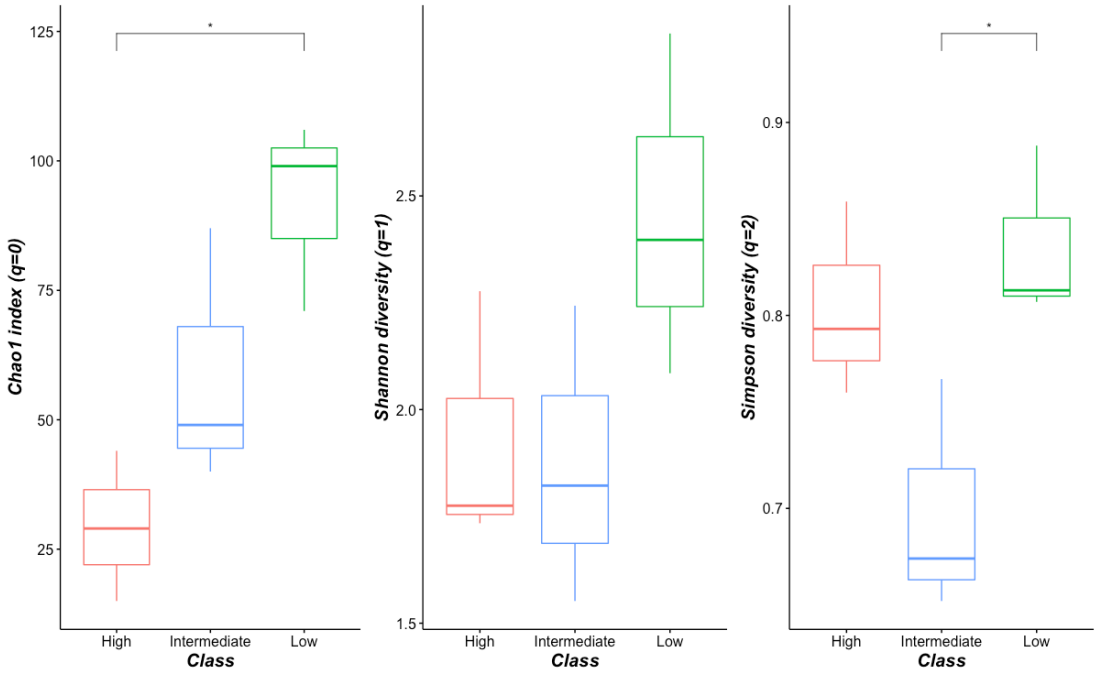


Figure 1.2 : **Species richness and diversity indices based on bacterial community. Effect of insect-densities on richness and diversity indices computed with the function iNext. The post hoc comparisons between the distinct levels using Tukey's 'Honest Significant Difference' method showed higher species richness for the 'Low insect density class' compared to the 'High insect density class. The Simpson estimator revealed a lower diversity for the 'Intermediate insect density class, whereas there was no significant difference when analyzing the Shannon estimator**

The bacterial microbiome was represented by six phyla, amongst which Proteobacteria (82%) and Actinobacteria (8%) were predominant (**Figure 1.3** and **Table 1.1**). More than 70% of the assigned genera were represented by Pantoea (35%), *Escherichia/Shigella* (14%), *Pseudomonas* (7%), *Chryseobacterium* (6%), *Propionibacterium* (6%) and *Streptococcus* (6%). The gut-associated fungal community was represented by the two phyla Ascomycota (88%) and Basidiomycota (12%). At least 80% of the identified genera included *Kabatiella* (43%), *Pyrenochaeta* (18%), *Preussia* (14%) and *Mycocentrospora* (6%). The core EAB microbiome consisted of taxa prevalent at 80% in all samples including seven ASVs, two fungi from the Kabatiella genus (the fungal core) and five bacteria from the Enterobacteriaceae family, among which one *Pantoea* genus and one *Escherichia* genus (the bacterial core) were found (**Table 1.2**).

1.5.2 Variations in the gut microbiome profile

For the bacterial community, the species richness assessed by the Chao1 estimator (q = 0) which is particularly sensitive to rare species showed significant differences across the three classes of insect population density with significant discrimination between the low-density and high-density classes (**Figure 1.2**). The greatest richness was seen in insects belonging to the low-density class. We observed no significant differences in Shannon's entropy (q = 1) among the three classes, whereas the Simpson

index (q = 2) known to be more sensitive to dominant species appeared significantly lower in the intermediate-density class. The α-diversity pattern of the fungal community was not consistent with insect densities (**Table Suppl. 1.2**, Supporting Information).

The hierarchical clustering of community profiles showed two significant groups for bacteria in contrast to fungi (**Figure Suppl. 1.1**, Supporting Information). Using bacterial community data, the measurements of β-diversity across the three classes by PCoA analysis based on weighted and unweighted Unifrac distances revealed sample clustering into the three density classes, with 52% and 43% of variations represented by the first two axes, respectively (**Figure 1.4**, section 'A' and section 'B'). For the weighted Unifrac distances, the first axis discriminates the low-density class on the bottom side from the intermediate-density and high-density classes on the top. In contrast, the second axis discriminates the intermediate-density class on the left from the two other classes on the right. With a 30% contribution cut-off, the first axis reflected seventy specific ASVs belonging to five phyla, the most abundant being the Proteobacteria (71%), represented by the Alphaproteobacteria (22%), the Betaproteobacteria (24%) and the Gammaproteobacteria (54%). The second axis had 64 ASVs belonging to six phyla, mostly Proteobacteria (74%), represented by the Gammaproteobacteria (80%), the Betaproteobacteria (13%) and the Alphaproteobacteria (7%). The microbiome composition was significantly related to the number of insects collected per tree, according to a PERMANOVA analysis of the partitioning of the three population density classes. The Wilcoxon rank-sum tests with the Bonferroni P value adjustment, revealed significant differences among the three classes with strong separation of bacterial communities between each class ($P < 0.001$, **Table 1.3**). The same analysis performed with OTUs showed the same profile, except for the unweighted Unifrac distances (**Figure 1.4**, section 'C'). In contrast to the bacterial community data, the results for the fungal micro- biota showed no significant differences among the classes (**Table Suppl. 1.3**, Supporting Information).

Examination of bacterial and fungal ASVs contributing to discriminating microbial community profiles amongst density classes led to the identification of eleven indicator ASVs (**Table 1.4**). In the low-density class, nine indicator ASVs were found, including eight Gammaproteobacteria and one Bacillus.

Table 1.1: **Bacterial and fungal communities associated with the microbiome of the EAB adult gut. The composition of bacterial and fungal communities associated with the microbiome of the A. planipennis adult gut among the samples showing some dominant phyla (%). Each class is represented by its percentage in the sample (relative abundancies of the phylum in each sample), issued from taxonomic assignation using Ribosomal Data Base (RDP) and the final ASVs abundance table. The last column indicates the means of relative values across all the samples. Unassigned phyla are ASVs which were not assigned during the taxonomic assignation step using the RDP.**

		Insect density class									
		Low			Intermediate			High			Mean
	Phylum	1	2	3	1	2	3	1	2	3	
Bacteria	Actinobacteria	14.4	0.0	0.7	0.3	6.8	4.2	9.7	0.0	32.2	7.6
	Bacteroidetes	0.0	24.0	3.3	0.7	2.2	0.0	1.5	2.7	3.0	4.2
	Deinococcus-Thermus	0.0	0.0	1.6	0.0	0.0	0.0	0.0	0.5	0.0	0.2
	Firmicutes	1.0	1.7	1.0	0.1	0.6	4.7	3.1	4.3	32.2	5.4
	Fusobacteria	0.0	0,1	0.5	0.0	0.0	0.0	0.0	0,0	0.0	0.1
	Proteobacteria	84.7	74.3	92.9	98.8	87.9	91.1	85.6	92.2	32.5	82.2
	Unassigned	0.0	0.0	0.0	0.0	2.5	0.0	0.0	0.3	0.0	0.3
Fungi	Ascomycota	95.1	98.1	79.5	53.9	69.8	100.0	100.0	100.0	98.5	88.3
	Basidiomycota	4.9	1.9	20.5	46.1	29.0	0.0	0.0	0.0	1.5	11.6
	Unassigned	0.0	0.0	0.0	0.0	1.3	0.0	0.0	0.0	0.0	0.1

Table 1.2: **The core microbiome of EAB adult corresponding to taxa prevalent at 80% across all samples with its abundances. "LOW" for the Low insect density class, "INT" for the Intermediate insect density class and "HIG" for high insect density class.**

	Class	Family	Genus	Low_1	Low_2	Low_3	Int_1	Int_2	Int_3	Hig_1	Hig_2	Hig_3
Asvb1	c_*γ*-proteobacteria	f_Enterobacteriaceae	g_Pantoea	2807	1774	4160	38 893	14 257	23 212	1	2	1
Asvb44	c_*γ*-proteobacteria	f_Enterobacteriaceae	g_Unassigned	518	459	629	573	2	2	3	2	0
Asvb6	c_*γ*-proteobacteria	f_Enterobacteriaceae	g_Unassigned	32	268	40	2	1	0	2195	13 797	1
Asvb7	c_*γ*-proteobacteria	f_Enterobacteriaceae	g_Escherichia/Shigella	17	65	914	284	6704	2718	943	1484	1827
Asvb46	c_*γ*-proteobacteria	f_Enterobacteriaceae	g_Unassigned	5	4	300	47	0	1	75	397	435
Asvf1	c_Dothideomycetes	f_Dothioraceae	g_Kabatiella	1643	43 615	31 471	5758	14 461	2	24 778	7682	12 385
Asvf18	c_Dothideomycetes	f_Dothioraceae	g_Kabatiella	9	133	93	16 425	22	0	61	23	30

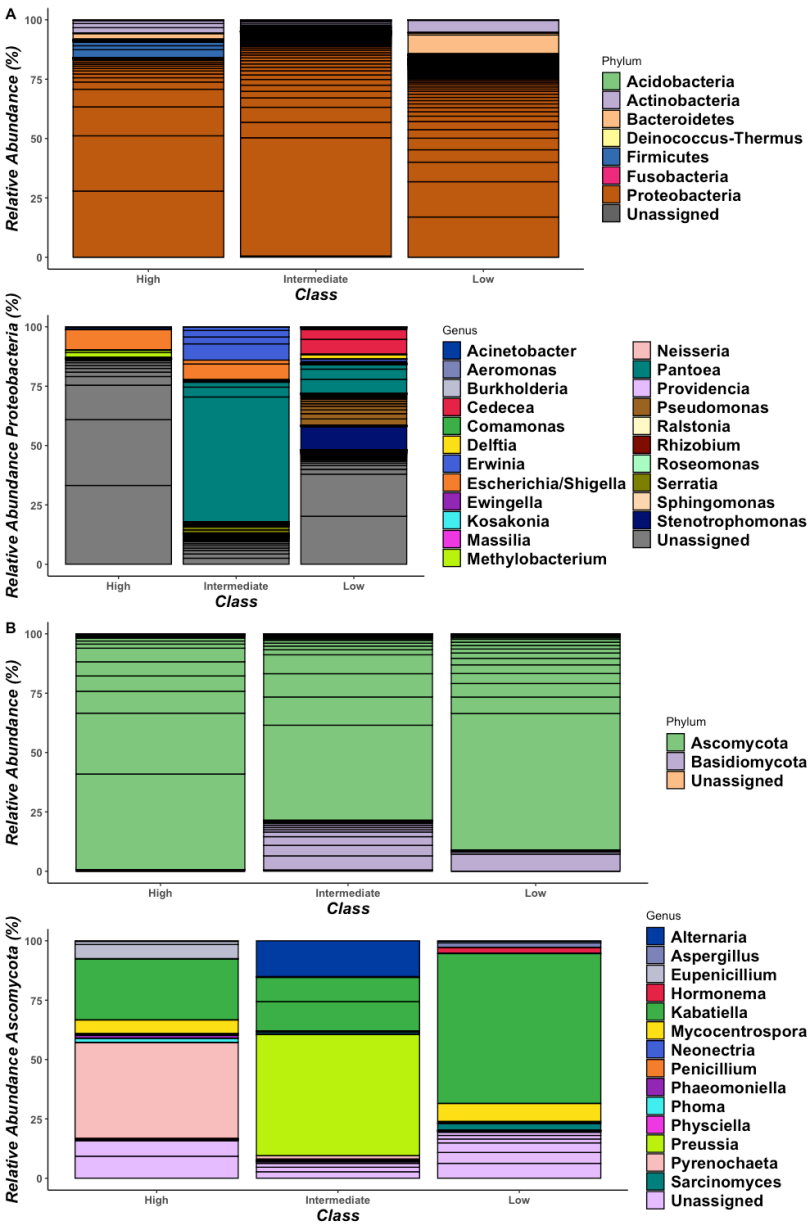


Figure 1.3 : **Composition of the bacterial and the fungal communities associated to the adult EAB gut. This representation is based on the number of sequences identified per ASV and the taxonomic assignation done with the RDP database. The proportion of the unassigned ASVs are also considered. The section 'A' shows the relative abundance of the bacterial community and the relative abundance of the predominant phylum (Proteobacteria). The section 'B' shows the relative abundance of the fungal community and the relative abundance of the predominant phylum (Ascomycota).**

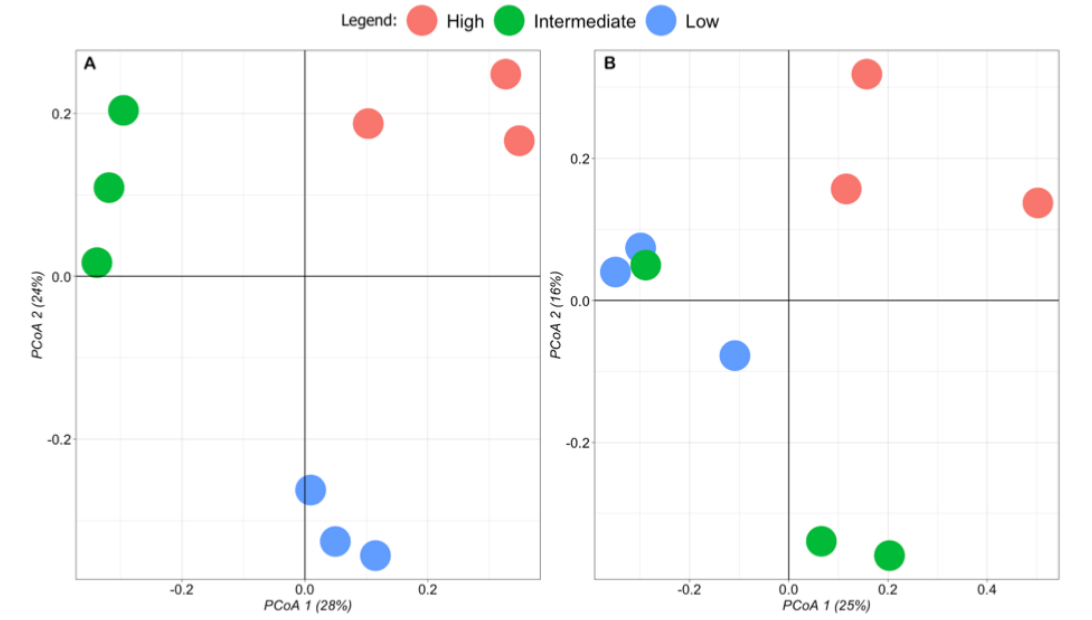
Table 1.3: The Wilcoxon rank-sum tests performed with ASV dataset, with a p value adjustment of Bonferroni, revealed a significant difference across the three classes with strong separation between each class for the bacterial community contrarily to the results observed with the fungal community.

Bacteria (weighted)		
	High	Intermediate
Intermediate	6.5e-07	–
Low	< 2e-16	4.2e-16
Bacteria (unweighted)		
	High	Intermediate
Intermediate	0.00073	–
Low	1e-08	0.05669

Table 1.4: The indicator ASVs characterizing taxa strongly correlated to the Low insects' density class, identified using the indicator species analysis procedure implemented in the 'indicspecies' package). Minimal significance level (alpha) of the indicator ASV was 0.05, tested against 9999 random permutations of samples among the three groups. The Association function: IndVal.g, Total number of ASVs: 258, 13 indicator ASVs have been found. The indicator ASVs identified by the sign "∇" were used for the primers' conception. The result was generated by using the Multipatt function of the indicspecies package, based on the species-site group association named Indval.g and with 9999 permutations. The highest Indval statistic indicate the accuracy of the taxa and the maximum is 1. The correlation was per- formed between ASV abundance values and the number of insects related to.

ASV_ID	Indicator value	Spearman correlation
Low insect density class		
ASVb187 (γ-proteobacteria)$^{\nabla}$	1.000*	− 0.8454**
ASVb192 (γ-proteobacteria)$^{\nabla}$	1.000*	− 0.8454**
ASVb170 (γ-proteobacteria)$^{\nabla}$	1.000*	− 0.8454**
ASVb3 (γ-proteobacteria)$^{\nabla}$	1.000*	− 0.8183**
ASVb18 (γ-proteobacteria)$^{\nabla}$	1.000*	− 0.8660**
ASVb2 (γ-proteobacteria)$^{\nabla}$	0.999*	− 0.7707*
ASVb37 (γ-proteobacteria)$^{\nabla}$	0.999*	− 0.8660**
ASVb74 (γ-proteobacteria)$^{\nabla}$	0.999*	− 0.8660**
ASVb62 (Bacilli)	0.997*	− 0.8660**
ASVb44 (γ-proteobacteria)	0.940*	− 0.8203***
Intermediate insect density class		
ASVb1 (γ-proteobacteria)	0.967*	− 0.4743 (NS)
High insect density class		
ASVf3 (Dothideomycetes)	0.984*	+ 0.9527**

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05, NS non-significant



C

ASV approach

	Df	Sum of Squares	Rsquared	F	p-value
Bacteria (weighted)					
Groups	2	1.4033	0.4928	2.9149	0.0039**
Residual	6	1.4443	0.5072		
Total	8	2.8477	1.00000		
Bacteria (unweighted)					
Groups	2	0.84512	0.32354	1.4349	0.0366*
Residual	6	1.76699	0.67646		
Total	8	2.61211	1.00000		

OTU approach

	Df	Sum of Squares	Rsquared	F	p-value
Bacteria (weighted)					
Groups	2	1.11954	0.53057	3.3907	0.0087**
Residual	6	0.99053	0.46943		
Total	8	2.11007	1.00000		
Bacteria (unweighted)					
Groups	2	0.28135	0.34816	1.6024	0.0622·
Residual	6	0.52675	0.65184		
Total	8	0.80809	1.00000		

Figure 1.4 : **Principal coordinate analysis (PCoA) performed on the bacterial community. The PCoA based on Unifrac distances unweighted (left, 'A') and weighted (right, 'B') distances of 16S rRNA sequences showed a significant variation between the structure of the bacterial communities depending on host tree infestation level (insect density level). The PERMANOVA test performed with bacterial community data, using the same distances and 9999 permutations revealed significant variations of the microbial community; more specifically changes were only on bacterial community. The percentages of variation explained by each axis are shown in parentheses. Post-hoc test showed significant differences between the three classes for bacterial community.**

The relevance of seven Gammaproteobacteria with an indica- tor value > 0.999 was challenged through an independent PCR assay (**Figure 1.5**). The results confirmed the specificity of the indicator taxa with amplicons for the three samples belonging to the low-density class. In contrast, no bands appeared for the sample belonging to the high-density class. That result confirmed that our results do not only rely on the compositional structure of ASV tables, which are often susceptible to spurious results. The Spearman correlation test showed a significant negative correlation between the identified indicator taxa and the proxy of insect density (**Tableau 1.4**). This suggests the importance of some Gammaproteobacteria in relation to EAB population size. The variations of the co-occurrences of the identified ASVs are shown in **Tableau Suppl. 1.5** (Supporting Information).

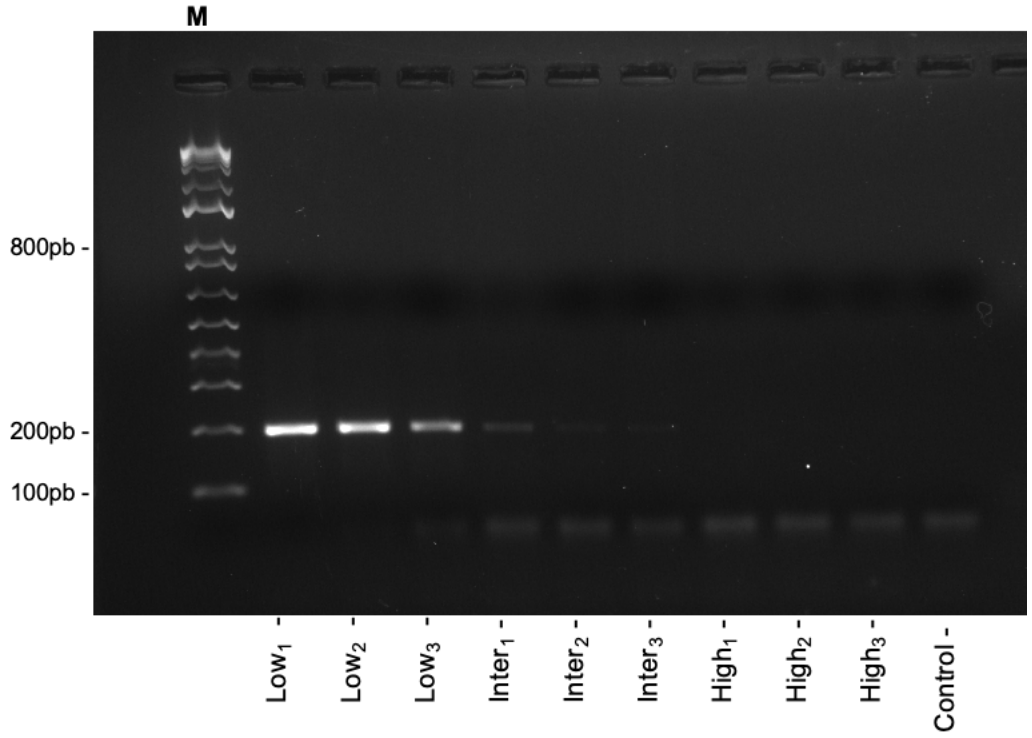


Figure 1.5 : Gel 2% agarose of the PCR amplification of six samples. The PCR assay was performed using template DNA from the different insect density classes to test the specificity of the indicator's species. Designed primers targeting indicator ASVs allowed to discriminate the 'High class' ($High_1$, $High_2$, $High_3$) from the two other classes. A negative control was included in the assay (Control). 'M' corresponds to the ladder.

1.6 Discussion

Considering Montreal as the epicenter of the EAB invasion, a northeast transect along the St. Lawrence River related to the variation of the number of captured insects was explained with the linear regression analysis. A previous report published by Siegert *et al.* (2010) showed that the probability of EAB colonization was affected by the distance from the epicenter, and in this study, insect density was categorized into a gradient defined by three classes: 'Low', 'Intermediate' and 'High'. We observed a significant negative regression between the number of insects captured and the distance from the highest insect density, which was located in the Montreal area (**Figure 1.2**). This study was designed to determine the composition of gut bacterial and fungal communities in adult EABs using high-throughput marker-gene sequencing and to explore the potential relationship between the number of adult EABs captured on a host tree and the taxonomic structure of the gut microbiota.

Insect gut microbiomes have been extensively studied, and it has been reported that most insect guts contain relatively few microbial species compared to mammalian guts (Engel & Moran, 2013). Our results showed a bacterial community of six phyla, mainly the Proteobacteria (82%) dominated by the Gammaproteobacteria class. The fungal community was less diverse with only two phyla, primarily the Ascomycota (88%) dominated by the Dothideomycetes. Many fungal taxa could not be identified at the

genus level with an 80% cut-off (42% unassigned). Some genera associated with the fungal community such as *Phoma, Aspergillus, Penicillium, Alternaria, Cryptococcus* and *Rhodotorula* have been recently reported in the gut of *Agrilus mali* Matsumura, an apple wood-boring beetle (Zhang *et al.*, 2018). Compared to the study of Vasanthakumar *et al.* (2008), based on 16S rRNA gene sequencing and culture approaches to characterize bacterial communities across life stages in EABs, the present study reveals more details about the bacterial community and shows for the first time, the profiles of the adult EAB gut bacterial and fungal communities based on high-throughput sequencing.

Using a conventional definition based on shared taxa, many studies have suggested the presence of a core microbiome for many families of Coleoptera orders such as the Tenebrionidae (Cambon *et al.*, 2018), Curculionidae (Mariño *et al.*, 2018), and Scolytidae (Berasategui *et al.*, 2016; Hernández-García *et al.*, 2017). According to the study of Vasanthakumar *et al.* (2008), the genera *Pantoea* and *Stenotrophomonas* occurred together in all samples of the EAB gut microbiome across the insects' life cycle. Another study suggested the co-occurrence of twelve specific bacterial genera in all adults (Bergeron, 2016). Based on a definition of the core microbiome as including taxa prevalent at 80% across all samples, our results revealed a core microbiome composed of seven taxa across all the samples amongst which were five members of the Enterobacteriaceae family, mainly *Pantoea* and *Escherichia*, and two members of the *Kabatiella* fungal genus. The co-occurrence of these persistent taxa suggests that they may play a crucial role for their hosts, particularly the *Pantoea* (ASVb1) and *Kabatiella* (ASVf1), which are the most abundant in terms of number of sequences. Some studies have reported *Kabatiella sp.* as an endophytic fungus in different plants or as a pathogen causing leaf diseases such as eyespot or northern anthracnose (Butin, 1992; Roy & Banerjee, 2018). *Pantoea sp.* may be able to synthesize four of the six enzymes responsible for plant-cell wall degradation (Bozorov *et al.*, 2019) and several species of *Pantoea* have pathogenic potential towards both plants and insects (Nadarasah & Stavrinides, 2014). Some have also been reported being responsible for epiphytic, endophytic and parasitic colonization of plants (Walterson & Stavrinides, 2015). According to Nadarasah and Stavrinides (2014), the constant cycling between *Pantoea* hosts in the environment can promote selective pressure for the acquisition of pathogenicity factors that can affect the host's ability to evolve resistance. Although the ubiquity of specific taxa such as *Kabatiella* and *Pantoea* genera suggests beneficial interactions between insect host and gut microbiota, most detected ASVs showed more complex distribution patterns, highlighting the potential influence of environment on the taxonomic composition of the EAB gut microbiome.

Among the broad array of factors defining the recruitment of bacteria and fungi to the host and the relative gut hospitability for different species, this study shows that the taxonomic structure of the adult EAB gut microbiome is correlated to population density. Both alpha and beta diversity of bacterial communities supported such a relationship. Among the Hill numbers examined in this study, the Chao1

index known to be most sensitive to rare ASV (q = 0) and the Simpson diversity as most sensitive to dominant ASV (q = 2) appeared sensitive to the insect population size. Interestingly, the intermediate-density class appeared as a transition stage where Simpson diversity reached its lowest level. This observation is supported by the distribution of the dominant member of the core microbiome namely, the ASVb1 assigned to *Pantoea sp*. that was the most abundant ASV in the intermediate-density class at 93% compared to 7% in low and 0.02% in high. Such a dominance of ASVb1 was accompanied by the absence of some taxa identified at the genus level such as *Pantoea* (ASVb14, 147, 151, 189 and 9), *Pseudomonas* (ASVb104 and 166), *Kosakonia* (ASVb107), *Cedecea* (ASVb15), *Chryseobacterium* (ASVb4), *Propionibacterium* (ASVb91), *Methylobacterium* (ASVb71) and *Streptococcus* (ASVb91). The relationship between gut microbial community structure and insect density was further supported by an independent PCR assay (**Figure 1.5**). The selected subset of indicator ASVs identified during PCR amplicon sequencing was not detected in the high-density class. The Spearman correlation revealed that the number of insects collected per tree was inversely correlated with each indicator ASV identified (Spearman < -0.8). Although this result suggests that some indicator ASVs were strongly associated with the low-density class, it is difficult to explain why those taxa, all belonging to the Gammaproteobacteria class, are indicator ASVs as this class was formed by pooling insects from different trees with low population size.

There are multiple drivers influencing the relationship between the taxonomic structure of the adult EAB gut microbiome and population density observed in this study, but plant phytochemistry is a potential factor to be considered. EAB adults feed on ash leaves throughout their adult life, eating up to 1 cm^2 of foliage per day (Wei *et al.*, 2007). This attack is expected to trigger defense mechanisms in the tree, including secondary metabolites and volatiles (Wielkopolan & Obrepalska-Steplowska, 2016). Stress induced by EAB attack is expected to be the highest in tree encompassing 'High-density class' because the insect alternate between chewing and resting, mostly staying on the same tree during their adult lifetime Wang *et al.* (2010). Production of secondary metabolites and volatiles could be a significant driver to selectively recruit microbial species possessing the enzymatic machinery to cope with defensive secondary plant metabolites, which frequently are designed to repel insects, deter feeding or be detrimental (Villari *et al.*, 2016). Another factor that might explain the relationship between the taxonomic structure of the adult EAB gut microbiome and population density observed in this study is the distance between surveyed trees. Trees in close proximity are exposed to more similar climate, soil substrates and air quality conditions than trees sampled across a broader distance. Therefore, the significant variation of the number of captured insects along the transect can make the geographical distance between surveyed trees a potential confounding variable to the density proxy. Pairwise dissimilarity of gut microbial communities from high-density and intermediate-density classes (each replicate was represented by a single tree) with the Mantel correlation (r = 0.62) suggests that there

must be an association with geographical distances. Nevertheless, the sample size is not enough to confirm that tendency (**Table S1.3**, Supporting Information). A more extensive survey, including more sites, would be necessary to assess the impact of geographical distance on the EAB's gut microbial community structure.

To conclude, our study presents an overview of the adult EAB gut's bacterial and fungal communities. The results suggest that the plasticity of the gut microbiome is related to population size, which raises two interesting questions: How do bacterial members interact in relation to insect population size? Could the observed taxonomic variations in the bacterial community be related to the host stress from herbivory? A more extensive survey including more sites and metadata (e.g. foliar chemistry) would be necessary to test the impact of such environmental factors on the gut microbial community structure of EAB.

Supplementary data are available at FEMSEC online.

1.7 Acknowledgements

We thank Philippe Labrie (Canadian National Resources, Quebec, Canada) and Narin Srei (INRS-Centre Armand-Frappier Santé Biotechnologie) for providing their technical assistance for the field sampling across the large area covered by this study.

This work was supported by a Research Funding Agreement (GDG-INRS-NRCAN) to Claude Guertin and Robert Lavallée. Judith Mogouong was financially supported by the Alexander Graham Bell Canada Scholarship (Natural Sciences and Engineering Research Council of Canada, [grant number 627371]), the Smirnoff fellowship and the INRS-Armand-Frappier Santé Biotechnologie fellowship.

1.8 Author contributions

J.M., P.C., R.L. and C.G. conceived and designed the experiments. J.M., P.C. and C.G. performed the experiments. J.M., P.C. and C.G. analyzed the data. J.M. and C.G. prepared the figures. P.C. and C.G. contributed reagents, materials and analysis tools. J.M., R.L. and C.G. performed the field sampling. J.M., P.C. and C.G. wrote the main manuscript text. All authors reviewed the manuscript.

Conflict of interest. We declare no conflict of interest.

1.9 Supplementary materials

Figure Suppl. 1.1: The linear regression of population density on distance showed a dispersion radiating from the epicenter in Montreal with the following equation: N = 4.44 - 0.06 *d

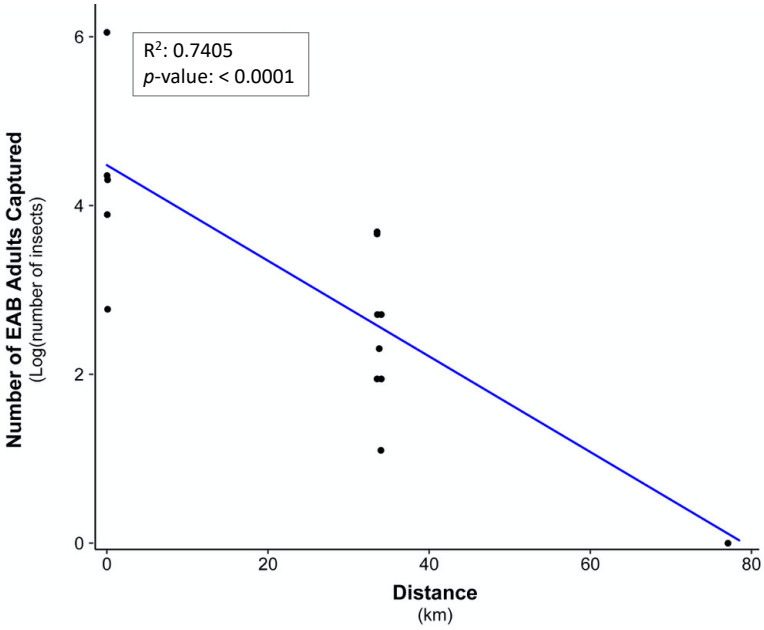


Figure Suppl. 1.2: The hierarchical clustering of community profiles showed two significant clusters for bacteria

[1] "HIG_3" "HIG_1" "HIG_2"

[2] "INT_2" "INT_3" "INT_1" "LOW_1" "LOW_2" "LOW_3"

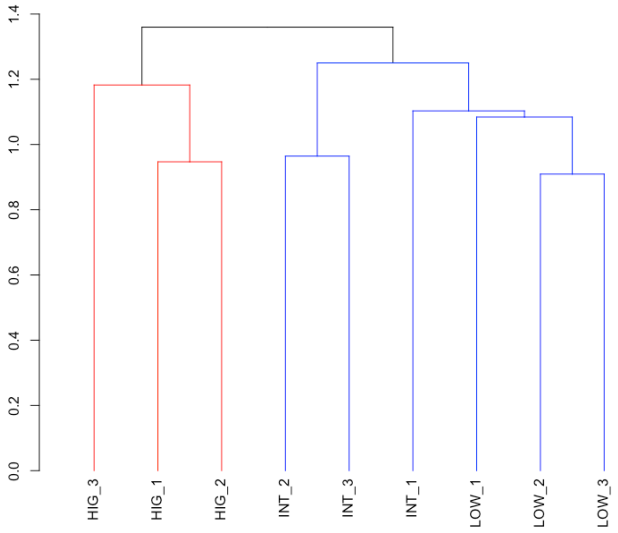


Table Suppl. 1.1: Bioinformatic output relative to the sequences processing

1. Bacteria sequences

Merged length distribution:
300 Min
438 Low quartile
438 Median
439 High quartile
482 Max
Totals before quality control :
591327 Pairs (591.3k)
483715 Merged (483.7k, 81.80%)
271699 Alignments with zero diffs (45.95%)
69851 Too many diffs (> 5) (11.81%)
36954 No alignment found (6.25%)
0 Alignment too short (< 16) (0.00%)
807 Merged too short (< 300)
0 Merged too long (> 500)
720 Staggered pairs (0.12%) merged & trimmed
62.82 Mean alignment length
437.18 Mean merged length
0.59 Mean fwd expected errors
1.06 Mean rev expected errors
0.81 Mean merged expected errors
OTU
Before trimming
453073 Reads
9 Samples
136 OTUs
After trimming 0.005%
452907 Reads
9 Samples
112 OTUs
ASV
Before trimming
453615 Reads
9 Samples
193 ASVs
After trim 0.005%
453530 Reads
9 Samples
187 ASVs

2. Fungi sequences

Merged length distribution:

250 Min

385 Low quartile

387 Median

389 High quartile

484 Max

Totals before quality control :

772146 Pairs (772.1k)

592908 Merged (592.9k, 76.79%)

316472 Alignments with zero diffs (40.99%)

147737 Too many diffs (> 5) (19.13%)

16955 No alignment found (2.20%)

0 Alignment too short (< 16) (0.00%)

14546 Merged too short (< 250)

0 Merged too long (> 500)

14856 Staggered pairs (1.92%) merged & trimmed

113.53 Mean alignment length

386.47 Mean merged length

0.48 Mean fwd expected errors

0.93 Mean rev expected errors

0.30 Mean merged expected errors

OTU

Before trimming

519230 Reads

9 Samples

107 OTUs

After trimming 0.005%

519057 Reads

9 Samples

88 OTUs

ASV

Before trimming

514858 Reads

9 Samples

73 OTUs

After trimming 0.005%

514813 Reads

9 Samples

71 OTUs

Table Suppl. 1.2: Species richness and diversity index performed in R software with the iNext function (iNEXT package)

	Index (se)	Low1	Low2	Low3	Int1	Int2	Int3	Hig1	Hig2	Hig3	Mean
Bacteria	Chao1 (9.574)	71	106	99	87	49	40	29	44	15	60
	Shannon (0.123)	2.085	2.397	2.880	2.243	1.552	1.822	2.277	1.775	1.734	2.085
	Simpson (0.023)	0.813	0.807	0.888	0.767	0.674	0.652	0.859	0.760	0.793	0.779
Fungi	Chao1 (3.245)	24	12	7	17	38	8	4	6	18	14
	Shannon (0.087)	0.856	0.277	0.453	0.724	0.85	0.001	0.498	0.729	0.799	0.576
	Simpson (0.241)	2.265	0.632	0.834	1.403	2.436	0.002	0.726	1.346	1.88	1.28

Table Suppl. 1.3. PERMANOVA output assessing differences between the fungal community of three population density classes using Adonis2 function in R (distance Bray-Curtis on ASV dataset), 9999 permutations were used and the classes were included as a random effect.

ASV approach

	Df	Sum of Square	Rsquared	Pseudo-F	*p*-value
Fungi					
Groups	2	0.48262	0.25836	1.0451	0.439
Residual	6	1.38538	0.74164		
Total	8	1.86800	1.00000		

OTU approach

	Df	Sum of Square	Rsquared	Pseudo-F	*p*-value
Fungi					
Groups	2	0.55087	0.32591	1.4504	0.0561
Residual	6	1.13939	0.67409		
Total	8	1.69152	1.00000		

Table Suppl. 1.4: The Mantel statistic test based on Spearman's rank correlation rho performed with 9999 permutations between the “Intermediate insects’ density population class” (n=3) and the “High insects’ density population class” (n=3)

Mantel statistic r	0.6214		
Significance	0.58333		
Upper quantiles of permutations (null model)			
90%	95%	97.5%	99%
0.819	0.846	0.850	0.863

Table Suppl. 1.5: Cooccurrences of the indicator ASVs across all the samples

ASV_ID	LOW_1	LOW_2	LOW_3	INT_1	INT_2	INT_3	HIG_1	HIG_2	HIG_3
Low insect density class									
ASVb187 (γ-proteobacteria)$^{\nabla}$	687	420	33	0	0	0	0	0	0
ASVb176 (γ-proteobacteria)$^{\nabla}$	8	25	1	0	0	0	0	0	0
ASVb3 (γ-proteobacteria)$^{\nabla}$	35744	304	69	5	0	2	3	18	0
ASVb18 (γ-proteobacteria)$^{\nabla}$	167	3965	8	0	1	0	0	0	0
ASVb2 (γ-proteobacteria)$^{\nabla}$	979	28379	11820	47	0	0	7	28	0
ASVb37 (γ-proteobacteria)$^{\nabla}$	383	19	1539	1	1	0	0	0	0
ASVb74 (γ-proteobacteria)$^{\nabla}$	163	514	199	5	0	0	0	0	0
ASVb62 (Bacilli)	837	17	70	7	0	0	0	0	0
ASVb44 (γ-proteobacteria)	518	459	629	573	2	2	3	2	0
Intermediate insect density class									
ASVb1 (γ-proteobacteria)	2807	1774	4160	38893	14257	23212	1	2	1
High insect density class									
ASVf3 (Dothideomycetes)	0	1	0	983	1376	5	32495	15502	22758

BIBLIOGRAPHIE (ARTICLE 1)

Aukema JE, Leung B, Kovacs K, Chivers C, Britton KO, Englin J, Frankel SJ, Haight RG, Holmes TP, Liebhold AM (2011) Economic impacts of non-native forest insects in the continental United States. *PLoS One* 6(9):e24587.

Berasategui A, Axelsson K, Nordlander G, Schmidt A, Borg-Karlson AK, Gershenzon J, Terenius O, Kaltenpoth M (2016) The gut microbiota of the pine weevil is similar across Europe and resembles that of other conifer-feeding beetles. *Molecular ecology* 25(16):4014-4031.

Bergeron A (2016) *Caractérisation de la communauté bactérienne associée à l'agrile du frêne, Agrilus planipennis Fairmaire (Coleoptera: Buprestidae)*. (Université du Québec, Institut National de la Recherche Scientifique).

Bonaventure G, VanDoorn A, Baldwin IT (2011) Herbivore-associated elicitors: FAC signaling and metabolism. *Trends in plant science* 16(6):294-299.

Bozorov TA, Rasulov BA, Zhang D (2019) Characterization of the gut microbiota of invasive Agrilus mali Matsumara (Coleoptera: Buprestidae) using high-throughput sequencing: uncovering plant cell-wall degrading bacteria. *Scientific Reports* 9(1):4923.

Broderick G (2004) Effect of low level monensin supplementation on the production of dairy cows fed alfalfa silage. *Journal of Dairy Science* 87(2):359-368.

Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR (2016) Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. *PLOS Biology* 14(11):e2000225.

Butin H (1992) Effect of endophytic fungi from oak (Quercus robur L.) on mortality of leaf inhabiting gall insects. *European Journal of Forest Pathology* 22(4):237-246.

Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal* 11(12):2639.

Cambon M, Ogier J-C, Lanois A, Ferdy J-B, Gaudriault S (2018) Changes in rearing conditions rapidly modify gut microbiota structure in Tenebrio molitor larvae. *bioRxiv* 10.1101/423178:423178.

Cappaert D, McCullough DG, Poland TM, Siegert NW (2005b) Emerald ash borer in North America: a research and regulatory challenge.

Chattopadhyay P, Banerjee G, Mukherjee S (2017) Recent trends of modern bacterial insecticides for pest control practice in integrated crop management system. *3 Biotech* 7(1):60.

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42(Database issue):D633-642.

Comeau AM, Li WKW, Tremblay J-É, Carmack EC, Lovejoy C (2011) Arctic Ocean Microbial Community Structure before and after the 2007 Record Sea Ice Minimum. *PLOS ONE* 6(11):e27492.

De Cáceres M & Jansen F (2016) Package 'indicspecies': relationship between species and groups of sites. R Package Version 1.7. 6.).

De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* 119(10):1674-1684.

Delalibera I, Vasanthakumar A, Klepzig KD, Raffa KF (2007) Composition of the bacterial community in the gut of the pine engraver, Ips pini (Say)(Coleoptera) colonizing red pine. Symbiosis. *Symbiosis* 43:97-104.

Douglas AE (2009) The microbial dimension in insect nutritional ecology. *Functional Ecology* 23(1):38-47.

Douglas AE (2015) Multiorganismal insects: diversity and function of resident microorganisms. *Annual review of entomology* 60:17-34.

Duan JJ, Bauer LS, Hansen JA, Abell KJ, Van Driesche R (2012) An improved method for monitoring parasitism and establishment of Oobius agrili (Hymenoptera: Encyrtidae), an egg parasitoid introduced for biological control of the emerald ash borer (Coleoptera: Buprestidae) in North America. *Biological Control* 60(3):255-261.

Durand A-A, Bergeron A, Constant P, Buffet J-P, Déziel E, Guertin C (2015) Surveying the endomicrobiome and ectomicrobiome of bark beetles: The case of Dendroctonus simplex. 5:17190.

Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv* :081257.

Edgar RC & Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 31(21):3476-3482.

Emerald ash Borer information network (2020) *Cooperative Emerald Ash Borer Project.* http://www.emeraldashborer.info/ (Consulté le March 2nd, 2020)

Engel P & Moran NA (2013) The gut microbiota of insects–diversity in structure and function. *FEMS microbiology reviews* 37(5):699-735.

Gandhi KJ & Herms DA (2010a) North American arthropods at risk due to widespread Fraxinus mortality caused by the alien emerald ash borer. *Biological Invasions* 12(6):1839-1846.

Haack RA, Jendek E, Liu H, Marchant KR, Petrice TR, Poland TM, Ye H (2002) The Emerald Ash Borer: A New Exotic Pest in North America.

Herms DA & McCullough DG (2014) Emerald ash borer invasion of North America: history, biology, ecology, impacts, and management. *Annu Rev Entomol* 59:13-30.

Herms DA, Stone AK, Chatfield JA (2004) Emerald ash borer: the beginning of the end of Ash in North America? *Special Circular-Ohio Agricultural Research and Development Center* :62-71.

Hernández-García JA, Briones-Roblero CI, Rivera-Orduña FN, Zúñiga G (2017) Revealing the gut bacteriome of Dendroctonus bark beetles (Curculionidae: Scolytinae): diversity, core members and co-evolutionary patterns. *Scientific Reports* 7(1):13864.

Howe GA & Schaller A (2008) Direct defenses in plants and their induction by wounding and insect herbivores. *Induced plant resistance to herbivory,* Springer. p 7-29.

Hsieh T, Ma K, Chao A (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods in Ecology and Evolution* 7(12):1451-1456.

Huang S & Zhang H (2013) The impact of environmental heterogeneity and life stage on the hindgut microbiota of Holotrichia parallela larvae (Coleoptera: Scarabaeidae). *PLoS One* 8(2):e57169.

Hughes CC, Johns RC, Sweeney JD (2014) A technical guide to installing beetle traps in the upper crown of trees. *Journal of the Acadian Entomological Society* 10.

Hurst GDD (2017) Extended genomes: symbiosis and evolution. *Interface Focus* 7(5):20170001.

Kovacs KF, Haight RG, McCullough DG, Mercader RJ, Siegert NW, Liebhold AM (2010) Cost of potential emerald ash borer damage in U.S. communities, 2009–2019. *Ecological Economics* 69(3):569-578.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* 33(7):1870-1874.

Lahti L, Shetty S, Salojarvi J (2017) Tools for microbiome analysis in R. .

Lyons DB, Caister C, De Groot P, Hamilton B, Marchant K, Scarr T, Turgeon J (2007) Survey guide for detection of emerald ash borer.

Mariño YA, Ospina OE, Verle Rodrigues JC, Bayman P (2018) High diversity and variability in the bacterial microbiota of the coffee berry borer (Coleoptera: Curculionidae), with emphasis on Wolbachia. *Journal of Applied Microbiology* 125(2):528-543.

Mascarin GM & Jaronski ST (2016) The production and uses of Beauveria bassiana as a microbial insecticide. *World Journal of Microbiology and Biotechnology* 32(11):177.

Medina R, Nachappa P, Tamborindeguy C (2011) Differences in bacterial diversity of host-associated populations of Phylloxera notabilis Pergande (Hemiptera: Phylloxeridae) in pecan and water hickory. *Journal of Evolutionary Biology* 24(4):761-771.

Montagna M, Chouaia B, Mazza G, Prosdocimi EM, Crotti E, Mereghetti V, Vacchini V, Giorgi A, De Biase A, Longo S (2015a) Effects of the diet on the microbiota of the red palm weevil (Coleoptera: Dryophthoridae). *PLoS One* 10(1):e0117439.

Montagna M, Gómez-Zurita J, Giorgi A, Epis S, Lozzia G, Bandi C (2015b) Metamicrobiomics in herbivore beetles of the genus Cryptocephalus (Chrysomelidae): toward the understanding of ecological determinants in insect symbiosis. *Insect Science* 22(3):340-352.

Morales-Jiménez J, Zúñiga G, Ramírez-Saad HC, Hernández-Rodríguez C (2012) Gut-associated bacteria throughout the life cycle of the bark beetle Dendroctonus rhizophagus Thomas and Bright (Curculionidae: Scolytinae) and their cellulolytic activities. *Microbial Ecology* 64(1):268-278.

Nadarasah G & Stavrinides J (2014) Quantitative evaluation of the host-colonizing capabilities of the enteric bacterium Pantoea using plant and insect hosts. *Microbiology* 160(3):602-615.

Nikoh N, Hosokawa T, Oshima K, Hattori M, Fukatsu T (2011) Reductive evolution of bacterial genome in insect gut environment. *Genome Biol Evol* 3:702-714.

Oksanen J, Kindt R, Legendre P, O'Hara B, Stevens MHH, Oksanen MJ, Suggests M (2007) The vegan package. *Community ecology package* 10:631-637.

Panteleev DY, Goryacheva II, Andrianov BV, Reznik NL, Lazebny OE, Kulikov AM (2007) The endosymbiotic bacterium Wolbachia enhances the nonspecific resistance to insect pathogens and alters behavior of Drosophila melanogaster. *Russian Journal of Genetics* 43(9):1066-1069.

Paracer S & Ahmadjian VS (1999) *Symbiosis : an introduction to biological associations.* Oxford University Press, New York, 2nd ed.

Poland TM, McCullough DG, dr Groot P, Grant G, MacDonald L, Cappaert DL (2005) Progress toward developing trapping techniques for the emerald ash borer. *In: Mastro, Victor; Reardon, Richard, comps. Proceedings of the emerals ash borer research and development meeting; 2004 October 5-6; Romulus, MI. FHTET 2004-15. Morgantown, WV: US Forest Service, Forest Health Technology Enterprise Team: 53-54.*

Popa V, Deziel E, Lavallee R, Bauce E, Guertin C (2012) The complex symbiotic relationships of bark beetles with microorganisms: a potential practical approach for biological control in forestry. *Pest Manag Sci* 68(7):963-975.

Priya NG, Ojha A, Kajla MK, Raj A, Rajagopal R (2012) Host plant induced variation in gut bacteria of Helicoverpa armigera. *PloS one* 7(1):e30768.

Rani A, Sharma A, Rajagopal R, Adak T, Bhatnagar RK (2009) Bacterial diversity analysis of larvae and adult midgut microflora using culture-dependent and culture-independent methods in lab-reared and field-collected Anopheles stephensi-an Asian malarial vector. *BMC microbiology* 9(1):96.

Roe AD, Rice AV, Coltman DW, Cooke JEK, Sperling FAH (2011) Comparative phylogeography, genetic differentiation and contrasting reproductive modes in three fungal symbionts of a multipartite bark beetle symbiosis. *Molecular Ecology* 20(3):584-600.

Rosenthal GA & Berenbaum MR (2012) *Herbivores: their interactions with secondary plant metabolites: ecological and evolutionary processes.* Academic Press,

Roy S & Banerjee D (2018) Diversity of Endophytes in Tropical Forests. *Endophytes of Forest Trees,* Springer. p 43-62.

Scott DW (2009) Sturges' rule. *Wiley Interdisciplinary Reviews: Computational Statistics* 1(3):303-306.

Siegert NW, McCullough D, Williams D, Fraser I, Poland T, Pierce S (2010) Dispersal of Agrilus planipennis (Coleoptera: Buprestidae) from discrete epicenters in two outlier sites. *Environmental Entomology* 39(2):253-265.

Snyder RL (1985) Hand calculating degree days. *Agricultural and forest meteorology* 35(1-4):353-358.

Starnes RL, Liu CL, Marrone PG (1993) History, use, and future of microbial insecticides. *American Entomologist* 39(2):83-91.

Steinhaus EA (1958) Crowding as a possible stress factor in insect disease. *Ecology* 39(3):503-514.

Team RC (2013) R: A language and environment for statistical computing.

Toju H, Tanabe AS, Yamamoto S, Sato H (2012) High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PloS one* 7(7):e40863.

Vasanthakumar A, Handelsman J, Schloss PD, Bauer LS, Raffa KF (2008) Gut microbiota of an invasive subcortical beetle, Agrilus planipennis Fairmaire, across various life stages. *Environmental Entomology* 37(5):1344-1353.

Venables W, Ripley B, Isbn S (2002) Statistics Complements to Modern Applied Statistics with S Fourth edition by.

Villari C, Herms DA, Whitehill JG, Cipollini D, Bonello P (2016) Progress and gaps in understanding mechanisms of ash tree resistance to emerald ash borer, a model for wood-boring insects that kill angiosperms. *New Phytologist* 209(1):63-79.

Walterson AM & Stavrinides J (2015) Pantoea: insights into a highly versatile and diverse genus within the Enterobacteriaceae. *FEMS Microbiology Reviews* 39(6):968-984.

Wang, Yang Z, Gould J, Zhang Y, Liu G, Liu E (2010) The biology and ecology of the emerald ash borer, Agrilus planipennis, in China. *J Insect Sci* 10:128.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology* 73(16):5261-5267.

Wei X, Wu Y, Reardon R, SUN TH, Lu M, SUN JH (2007) Biology and damage traits of emerald ash borer (*Agrilus planipennis* Fairmaire) in China. *Insect Science* 14(5):367-373.

Whitaker D & Christman M (2014) Package 'Clustsig'R Package Version 1.1 2014. *Disponible online en: https://cran. rproject. org/web/packages/clustsig/clustsig. pdf*.

Wielkopolan B & Obrepalska-Steplowska A (2016) Three-way interaction among plants, bacteria, and coleopteran insects. *Planta* 244(2):313-332.

Zhang Z, Jiao S, Li X, Li M (2018) Bacterial and fungal gut communities of Agrilus mali at different developmental stages and fed different diets. *Scientific Reports* 8(1):15634.

Zhao L, Lu M, Niu H, Fang G, Zhang S, Sun J (2013) A native fungal symbiont facilitates the prevalence and development of an invasive pathogen–native vector symbiosis. *Ecology* 94(12):2817-2826.

Zindel R, Gottlieb Y, Aebi A (2011) Arthropod symbioses: a neglected parameter in pest- and disease-control programmes. *Journal of Applied Ecology* 48(4):864-872.

2 THE PHYLLOSPHERE MICROBIOME OF HOST TREES CONTRIBUTES MORE THAN LEAF PHYTOCHEMICALS TO VARIATION IN THE AGRILUS PLANIPENNIS FAIRMAIRE GUT MICROBIOME STRUCTURE

Titre de l'article en français Le microbiome de la phyllosphère des arbres hôtes contributes beaucoup plus à la variation de la structure du microbiome associé au tract intestinal de l'insecte que la phytochime des feuilles

Auteurs :

Judith Mogouong[1], Philippe Constant[1], Pierre Legendre[2], and Claude Guertin[1]✉

Affiliations professionnelles

[1]Institut National de la Recherche Scientifique, Centre Armand-Frappier Santé Biotechnologie, Laval, QC, Canada, H7V 1B7

[2]Département de sciences biologiques, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal, Québec H3C 3J7, Canada

✉ Auteur correspondant

Titre de la revue ou de l'ouvrage :

Scientific Reports

Soumis le 27 janvier 2021, accepté le 20 juillet 2021, publié le 05 aout 2021

Author Contributions

J.M., P.C, and C.G. conceived and designed the experiments. J.M. and C.G. performed the experiments and field works. J.M., P.C., P.L., and C.G. contributed to data and statistical analyses. J.M. and C.G. prepared the figures. P.C. and C.G. contributed reagents, materials, and analysis tools. J.M. wrote the main manuscript and all the authors reviewed it.

Additional Information

Competing interests: The authors declare no competing interests.

Lien entre l'article et le suivant :

Les résultats précédents suggèrent que la plasticité du microbiome intestinal est liée à la densité de la population, ce qui soulève deux questions intéressantes : Comment les membres bactériens interagissent-ils face à certains facteurs environnementaux, comme la densité des insectes, qui peut être liée au stress de l'arbre hôte ? Est-ce que ces variations sont liées aux changements qui se produisent dans les arbres hôtes lorsqu'ils font face à des attaques par des herbivores comme les agriles du frêne adultes, notamment les changements potentiels au niveau de la phytochimie des feuilles ? Dans cet article, nous avons cherché à comprendre si l'effet des propriétés des arbres hôtes pourraient avoir un effet significatif sur la structure des communautés bactériennes et fongiques associées au tract intestinal de l'insecte. Cela permettra d'évaluer l'importance des propriétés des arbres hôtes qui pourrait être important dans la compréhension du processus d'établissement de l'agrile du frêne en Amérique du Nord.

2.1 Résumé en français

La composition du microbiome des organismes vivants est étroitement liée aux fonctions essentielles qui déterminent l'aptitude de l'hôte à prospérer et à s'adapter à un écosystème particulier. Bien que de multiples facteurs, dont le stade de développement, le régime alimentaire et la coévolution hôte-microbe, aient été rapportés capables d'entraîner des changements dans la composition des structures du microbiome, très peu de tentatives ont été faites pour démêler leurs diverses contributions dans une approche globale. Ici, nous utilisons comme modèle l'agrile du frêne (*Agrilus planipennis* Fairmaire), un ravageur herbivore et une menace réelle pour les espèces de frênes d'Amérique du Nord, pour d'explorer les réponses du microbiome intestinal de l'agrile du frêne adulte confronté aux propriétés foliaires du frêne et d'identifier les prédicteurs potentiels des variations de la structure de ces communautés microbiennes. Les contributions relatives des propriétés spécifiques des arbres hôtes, à savoir les communautés bactériennes et fongiques sur les feuilles, la composition phytochimique et les coordonnées géographiques des sites d'échantillonnage, à la communauté microbienne intestinale de l'agrile du frêne ont été examinées par des analyses canoniques. La composition du microbiome de la phyllosphère semble être un puissant prédicteur de la structure de la communauté microbienne dans les intestins de l'agrile du frêne, expliquant respectivement 53 et 48 % de la variation des champignons et des bactéries. Cette étude suggère une covariation potentielle des microorganismes associés aux sources alimentaires et au microbiome intestinal des insectes.

2.2 Abstract

The microbiome composition of living organisms is closely linked to essential functions determining the fitness of the host for thriving and adapting to a particular ecosystem. Although multiple factors, including the developmental stage, the diet, and host-microbe coevolution have been reported to drive compositional changes in the microbiome structures, very few attempts have been made to disentangle their various contributions in a global approach. Here, we focus on the emerald ash borer (EAB), an herbivorous pest and a real threat to North American ash tree species, to explore the responses of the adult EAB gut microbiome to ash leaf properties, and to identify potential predictors of EAB microbial variations. The relative contributions of specific host plant properties, namely bacterial and fungal communities on leaves, phytochemical composition, and the geographical coordinates of the sampling sites, to the EAB gut microbial community was examined by canonical analyses. The composition of the phyllosphere microbiome appeared to be a strong predictor of the microbial community structure in EAB guts, explaining 53 and 48% of the variation in fungi and bacteria, respectively. This study suggests a potential covariation of the microorganisms associated with food sources and the insect gut microbiome.

2.3 Introduction

The study of environmental microbial communities may contribute to a deeper understanding of an ecosystem, particularly where their interactions with abiotic and biotic factors are considered. Living organisms are colonized by various microorganisms, encompassing bacteria, fungi, archaea, and protozoa (Feldhaar, 2011) that significantly participate in the host's essential physiological functions. The combination of compositional information about the microorganisms inhabiting a host and their genomes is defined as the host microbiome. Insects are commonly used as model systems for microbiome studies, and the microorganisms colonizing them are considered primarily symbiotic as they are closely involved in the performance of their host's essential ecological functions as well as their biological functions and fitness (Popa *et al.*, 2012). They are considered model systems because their gut has a lower diversity microbiome than mammalian systems, allowing cost-effective and time-efficient studies to investigate the complexity of microbial interactions (Qadri *et al.*, 2020). There is a growing body of evidence that insect gut microbiomes are shaped by variables such as the developmental stage

(Vasanthakumar *et al.*, 2008; Zhang *et al.*, 2018), diet (Colman *et al.*, 2012; Franzini *et al.*, 2016; Kim *et al.*, 2017), environment (Ferguson *et al.*, 2018), plant defense mechanisms (Mason *et al.*, 2015), and even the insect population density (Mogouong *et al.*, 2020). Several research efforts investigated the relationship between pest management and the insect microbiome, some suggesting microbiome manipulations (Qadri *et al.*, 2020). One approach example based on replacing the primary symbiont **Buchnera** with a specific genotype by microinjection in the pea aphid could alter the thermal tolerance of the insect (Moran & Yun, 2015). The potential contributions of the insect microbiome to diverse survival processes, including those related to its invasiveness traits, need to be addressed for efficient pest management approaches. Although studies of microbiome dynamics and plasticity have been increasing in number, most of them suffer from limitations including lack of a holistic approach for elucidating the mechanisms of complex ecological processes. One of the most frequently used analyses to investigate hypotheses related to the microbiome is redundancy analysis (RDA), a form of canonical analysis that is part of the regression modelling family. RDA is well suited for investigating variations in a response data matrix by one or several explanatory matrices. In addition, partial RDA allows researchers to identify redundancies among sets of explanatory variables. In this study, redundancy analysis is the computational workhorse in the variation partitioning method proposed by Borcard *et al.* (1992) and Peres-Neto *et al.* (2006) to determine the independent and joint contributions of multiple explanatory datasets and their redundancy in explaining host-microbiome variations. This approach is expected to improve understanding of the relationship between the host microbiome and the environment and help identify environmental predictors that explain the structure and composition of those microbial communities.

Among the Insecta, some species are considered threats to an ecosystem when they cause severe plant damage, especially to forest trees like the American ash. The emerald ash borer (EAB), Agrilus planipennis Fairmaire, is a holometabolous insect reported to cause significant environmental and economic damage to several Fraxinus species in North America (Aukema *et al.*, 2011; Cappaert *et al.*, 2005a; Kovacs *et al.*, 2010). After its introduction, EAB developed into an invasive pest in the highly urbanized area of Detroit, Michigan, USA (Poland & McCullough, 2006). As the life traits of EAB make the early detection very difficult, it continues to spread across the continent and is responsible for millions of ash tree's death (Herms & McCullough, 2014). That insect colonizes health and stressed ash trees and has been reported to be a disruption cause for natural processes threatening some native ash trees and, consequently, the forest diversity (McCullough, 2020). Moreover, several studies have shown that ash mortality can impact the hydrology in wetlands and harm insect species or fauna reported to be ash dependant (Gandhi

& Herms, 2010a; Slesak *et al.*, 2014). Despite all the control strategies implemented, including biological control and the introduction of natural enemies that contribute to slow the spread of EAB, it is almost impossible to stop its progression. During the host establishment, herbivorous insects may have to cope with defensive plant responses, mainly the production of molecules harmful to insects and the intensification of the lignification process (Howe & Schaller, 2008; Stam *et al.*, 2014; Wielkopolan & Obrepalska-Steplowska, 2016), and with microorganisms inhabiting the leaves of the host tree. Insects harbour many different biotopes in their bodies, but the gut is the most favourable one for colonization by microorganisms (Douglas, 2015). The gut is largely a protected environment for microorganisms, but it can present them with some adverse conditions, such as harmful phytochemicals and other factors shaping microbial community structure. Similarly, the phyllosphere, represented by the aerial part of the leaves (including endophytes and epiphytes) (Vorholt, 2012), can be associated with some microorganisms, which may help the plant facing pathogens. Although many studies reported on the effects of insect gut-associated microorganisms on leaf defences (Shikano *et al.*, 2017), the relationship between environmental microorganisms and the insect gut microbiota has received little attention. As the microbiome may contribute to the insect invasiveness processes, a better comprehension of its plasticity mechanisms based on a holistic approach should open new research avenues valuable for pest management. Since recent decades, that research field is getting more attention with some strategies targeting microbiome manipulation approaches that could alter the pest traits (Qadri *et al.*, 2020). Our previous works (Mogouong *et al.*, 2020) showed that some changes in the taxonomical structure of the bacterial community associated with the adult EAB gut could be related to the level of infestation of the host tree, suggesting that there may be a bipartite relationship between the “adult EAB gut microbiome” and its “host tree properties including their physiological traits”. This study was designed to test the hypothesis that variation in the insect gut microbiome can be attributed to host tree leaf phytochemicals, the phyllosphere microbiome, and geographical location. A field survey was conducted to address three complementary questions: (1) How are the microbial communities structured in the adult EAB gut and in the phyllosphere biotope (similarities and differences)? (2) Could the ash leaf microbiome and the phytochemical profile be predictors of the bacterial and fungal communities associated with adult EAB gut? And (3), to what extent do ash leaf microbiome and phytochemicals explain variations in the adult EAB gut microbiome? We compared the microbial communities of the insects and the host trees and performed a series of redundancy analyses to identify potential predictors among host tree leaf phytochemicals, phyllosphere microbiome, and geographic location for variations in the EAB gut microbiome. Selected predictors were then used to partition the variations observed within the

adult EAB gut microbial communities. The individual and redundant contributions of explanatory matrices were computed to model the observed variations in EAB gut microbiome structure.

2.4 Results

On the 18 selected trees, the number of EABs collected per tree ranged from two to 100 adults (**Figure 2.1)**. According to the principal component analysis (PCA), the most contributing phytochemicals to the variability of the sampling sites were cellulose, acid fibre, sucrose, and total non-structural carbohydrate (**Figure Suppl. S2.1**).

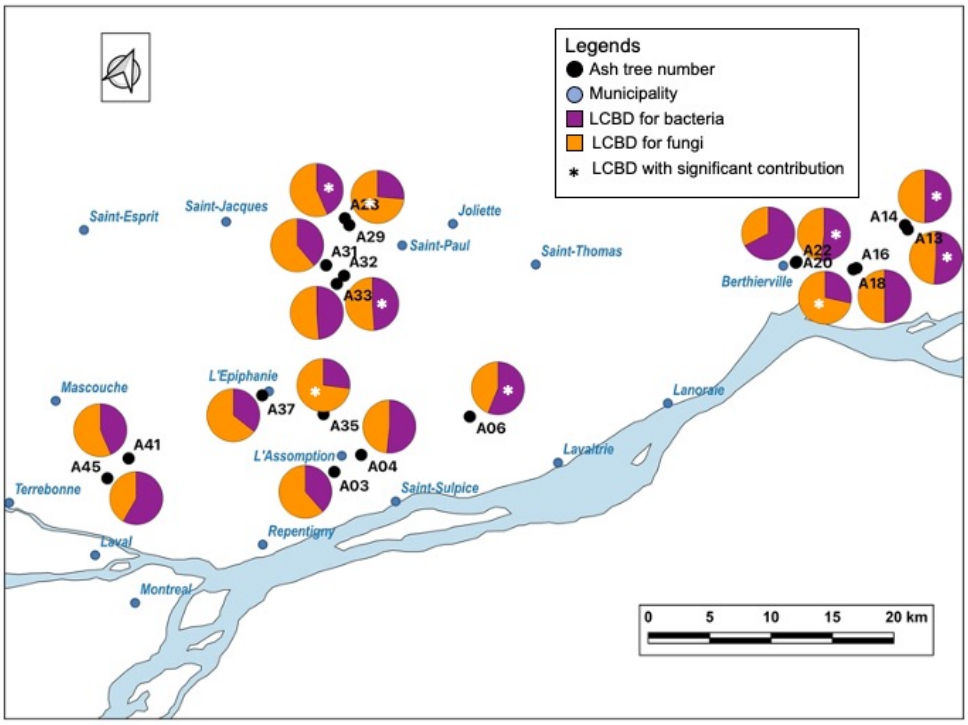


Figure 2.1 : Map representing sampling sites including the LCBD value per site. For each site LCBD values of bacteria (purple) and fungi (orange) are represented in percentages illustrated in a pie chart. The white star on the colour indicates the significant contribution of the corresponding community to the local β-diversity. Thus, the bacterial community was found significantly contributing to the local β-diversity in six sites (A06, A13, A14, A22, A28, and A32), Holm-corrected LCBD p values, whereas the fungal community was found significantly contributing to the local β-diversity in three sites (A18, A29, and A35), Holm-corrected LCBD p values.

Proteobacteria, Bacteriodetes, and Actinobacteria dominated the bacterial communities of the phyllosphere and EAB gut. A total of 186 amplicon sequence variants (ASVs) were shared by the two biotopes, 206 ASVs solely detected in the guts and 10 ASVs solely found in the leaves (**Figure 2.2 A and B).** Ascomycota was the main representative of fungal communities in the phyllosphere and the gut. The distribution of 111 ASVs encompassed both biotopes, whereas 31 ASVs were solely detected in the EAB gut and 62 ASVs in the leaves (**Figure 2.2 C and D**).

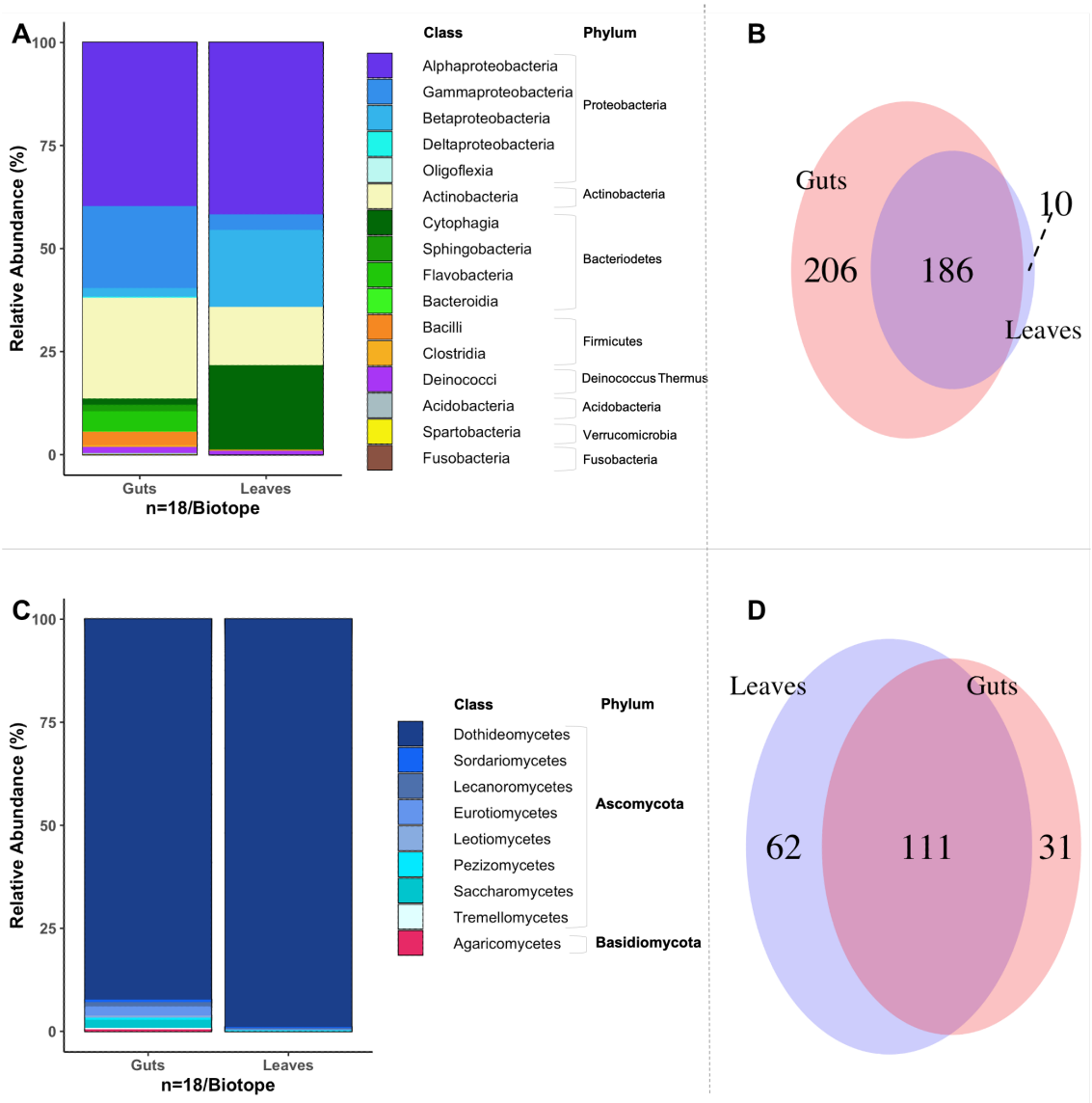


Figure 2.2 : Taxonomic profiles of bacterial (A) and fungal (C) communities associated with the adult EAB gut and those associated with the leaves of the host trees. The taxonomic profile is based on the presence/absence of ASVs in each biotope. The right part of the figure shows the number of ASVs shared and unshared between the two habitats for bacteria (B) and fungi (D).

Alpha- and beta-diversity of leaves and guts microbiomes were compared and related to ash leaf phytochemicals to examine linkages between foliar biotic and abiotic features and the EAB microbiome. The species richness of the guts bacterial community was similar to that in leaves, with means of 109 ± 87 and 108 ± 62, respectively. Targeted α-diversity indices showed contrasting patterns in the two biotopes. There was no significant difference in diversity between the two biotopes for the Simpson diversity (guts: 0.83 ± 0.21, leaves: 0.93 ± 0.07) and for phylogenetic diversity (guts: 8.80 ± 4.82, leaves: 10.61 ± 3.01), whereas a lower Shannon diversity index was observed in the guts (2.82 ± 1.18) compared to leaves (3.53 ± 0.57) (**Figure 2.3 A**). The guts environment exerted a stronger filtering effect on fungal communities with lower species richness (28 ± 14) than the leaves (110 ± 37) and a constrained phylogenetic diversity in the gut (15.6 ± 2.4) compared to the leaves (22.81 ± 3.16). These responses were mostly driven by rare taxa because neither the Shannon diversity index (guts: 1.35 ± 0.65, leaves: 1.78 ± 0.56) nor the Simpson diversity index (guts: 0.56 ± 0.25, leaves: 0.63 ± 0.16) differed between the two biotopes (**Figure 2.3 B).** Alpha diversity parameters of the bacterial community associated with the guts were not related to phytochemical variables, but a significant relationship was found between fungal species richness in the guts and the cellulose content ($F = 5.1629$, $p = 0.0382$).

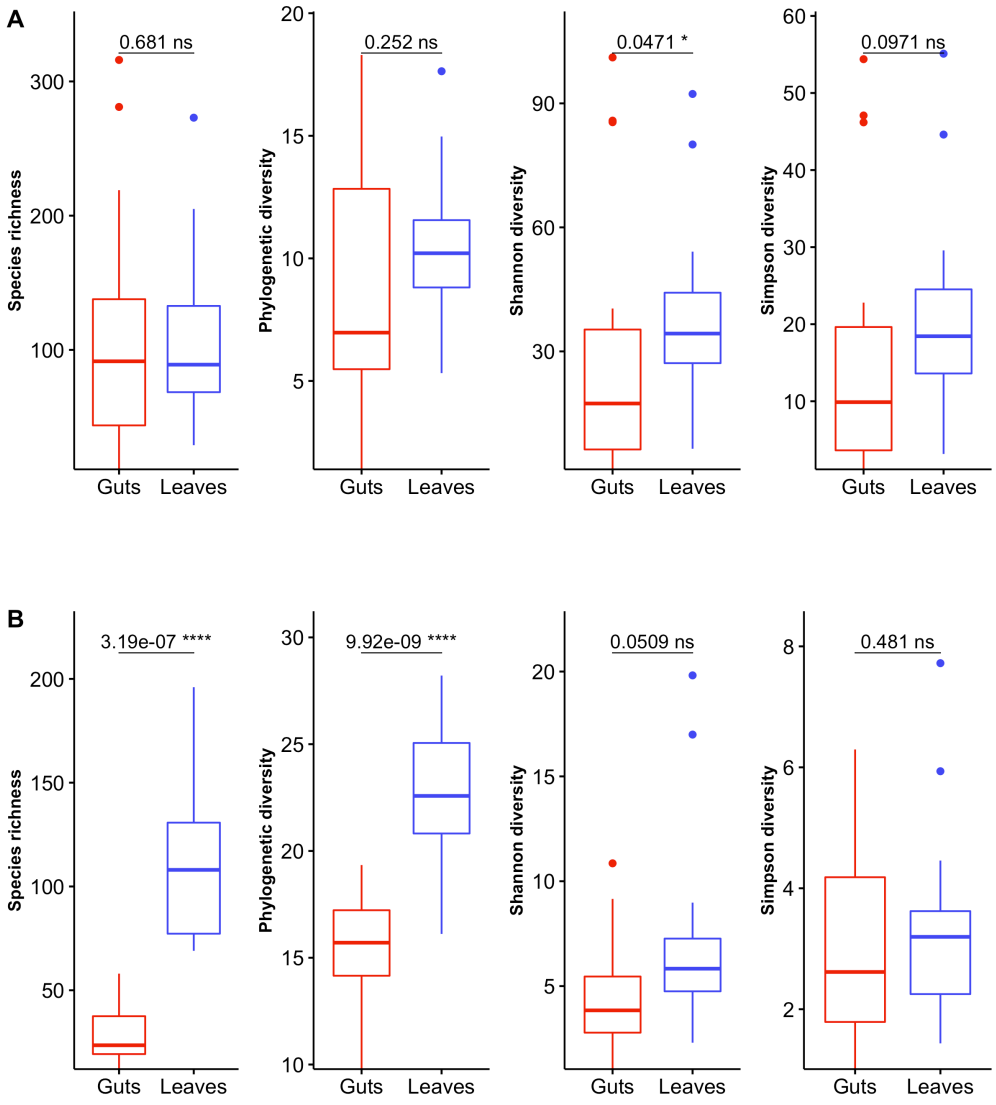


Figure 2.2 : Species richness and diversity indices in the insect gut and on the leaves computed from raw data for communities of bacteria (A) and fungi (B). Values at the top of the panels: Wilcoxon signed-rank test statistics and significance: Significance codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 'NS' > 0.05.**

The EAB guts and ash leaves displayed contrasting bacterial (F = 8.94, adj. R2 = 0.18, p< 0.001) and fungal (F = 5.79, adj. R2 = 0.12, p< 0.001) community profiles, according to a db-RDA (**Figure Suppl. 2.2**). This dissimilarity between microbial profiles was supported by the occurrence of indicator ASVs (**Table 2.1)**. Indicator ASVs for the guts were represented by 12 bacterial ASVs affiliated with Proteobacteria, while six bacteria (Proteobacteria and Bacteriodetes) and 14 fungi, mostly Dothideomycetes, were indicator ASVs for the leaves. Total variance in the microbial

community between the gut and the leaves appeared higher for bacteria (0.77) than for fungi (0.55). Indeed, the bacterial community revealed more species (ASVs) with contributions to the β-diversity (SCBD) greater than the mean SCBD found across all the samples compared to the fungal community, with respectively 82 ASVs and 26 ASVs (data not shown). An overview of the local contribution to the ß-diversity (LCBD) revealed greater values for the bacterial community in the insect gut (**Figure Suppl. S2.3).** More specifically, the bacterial community showed a higher LCBD mean value compared to the fungal community with respectively six sites (A06, A13, A14, A22, A28 and A32) and three sites (A18, A29, A35) having significant Holm-corrected LCBD p-values (**Figure 2.1**). Neither the LCBD nor the SCBD profiles were related to leaf phytochemicals based on the selection of explanatory variables in linear regression (data not shown). Similarly, the RDA constraining variation of the gut microbiome from the phytochemical profile was not significant. This decoupling between chemical and microbial profiles precluded consideration of phytochemicals in subsequent variation partitioning analyses.

Table 2.1 **The indicator ASVs characterizing taxa strongly correlated to each biotope, identified using the indicator species analysis procedure implemented in the 'indicspecies' package. The result was generated by using the multipatt() function of the indicspecies package, based on the species-site group association named Indval.g. The minimal significance alpha=0.001 and 9999 permutations of samples among the two biotopes. The highest Indval statistic indicates the accuracy of the taxa and the maximum is 1.**

ASVs[1]	Indval statistic	Phylum	Class	Order	Genus
Adult EAB Gut					
Asvb11	0.913 ***	Bacteroidetes	Flavobacteria	Flavobacteriales	*Flavobacterium*
Asvb24	0.912 ***	Proteobacteria	Gammaproteobacteria	Enterobacteriales	*Serratia*
Asvb14	0.882 ***	Proteobacteria	Alphaproteobacteria	Caulobacterales	*Caulobacter*
Asvb28	0.881 ***	Proteobacteria	Alphaproteobacteria	Rhizobiales	*Bradyrhizobium*
Asvb47	0.878 ***	Proteobacteria	Gammaproteobacteria	Pseudomonadales	*Pseudomonas*
Asvb20	0.849 ***	Proteobacteria	Gammaproteobacteria	Alteromonadales	*Alishewanella*
Asvb63	0.816 ***	Proteobacteria	Gammaproteobacteria	Xanthomonadales	*Luteibacter*
Asvb31	0.782 ***	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Unassigned
Asvb73	0.782 ***	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Unassigned
Asvb25	0.782 ***	Proteobacteria	Alphaproteobacteria	Caulobacterales	Unassigned
Asvb64	0.745 ***	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Unassigned
Asvb278	0.707 ***	Proteobacteria	Gammaproteobacteria	Pseudomonadales	*Pseudomonas*
Ash Leaves					
Asvb874	0.999 ***	Bacteroidetes	Cytophagia	Cytophagales	*Hymenobacter*
Asvf1	0.957 ***	Ascomycota	Dothideomycetes	Pleosporales	Unassigned
Asvf39	0.942 ***	Ascomycota	Unassigned	Unassigned	Unassigned
Asvb776	0.913 ***	Bacteroidetes	Cytophagia	Cytophagales	*Hymenobacter*
Asvb625	0.913 ***	Bacteroidetes	Cytophagia	Cytophagales	*Hymenobacter*
Asvf3	0.913 ***	Ascomycota	Dothideomycetes	Pleosporales	Unassigned
Asvf11	0.890 ***	Ascomycota	Dothideomycetes	Pleosporales	Unassigned
Asvf23	0.875 ***	Ascomycota	Dothideomycetes	Myriangiales	Unassigned
Asvf7	0.862 ***	Ascomycota	Dothideomycetes	Pleosporales	*Pyrenochaeta*
Asvf48	0.849 ***	Ascomycota	Dothideomycetes	Dothideales	Unassigned
Asvb56	0.845 ***	Proteobacteria	Alphaproteobacteria	Sphingomonadales	*Sphingomonas*
Asvb68	0.843 ***	Proteobacteria	Alphaproteobacteria	Sphingomonadales	*Sphingomonas*
Asvf89	0.821 ***	Ascomycota	Dothideomycetes	Capnodiales	*Mycosphaerella*
Asvf86	0.816 ***	Ascomycota	Dothideomycetes	Dothideales	Unassigned
Asvf46	0.813 ***	Ascomycota	Dothideomycetes	Pleosporales	Unassigned
Asvf159	0.782 ***	Ascomycota	Dothideomycetes	Pleosporales	Unassigned
Asvf215	0.779 ***	Ascomycota	Dothideomycetes	Capnodiales	*Mycosphaerella*
Asvf346	0.751 ***	Ascomycota	Dothideomycetes	Pleosporales	*Lewia*
Asvb903	0.707 ***	Proteobacteria	Betaproteobacteria	Burkholderiales	*Massilia*
Asvf141	0.707 ***	Ascomycota	Unassigned	Unassigned	Unassigned

[1] Asvb, ASV assigned to bacteria; Asvf, ASV related to fungi.

Variation in gut microbes: partitioning of the observed variation

The variation in the composition of the EAB gut microbiome was related to the leaf phyllosphere microbiome and the sampling site location. The variables that contributed the most to the changes in the EAB microbiome were selected in two partial RDAs. Six bacterial ASVs: ASVb105 (Gammaproteobacteria), ASVb869 (Gammaproteobacteria), ASVb78 (Actinobacteria), ASVb299 (Alphaproteobacteria), ASVb205 (Alphaproteobacteria), and ASVb5 (Actinobacteria), and six fungal ASVs from the Ascomycota, ASVf26, ASVf145, ASVf235, ASVf90, ASVf263, and ASVf173, associated with the leaves accounted for most of the variation in the composition of the gut bacterial community (**Table 2.2).** Two significant dbMEM, namely MEM6 ($F = 2.12$, adj. $R^2 = 0.06$, $p = 0.004$) and MEM7 ($F = 1.81$, adj. $R^2 = 0.11$, $p = 0.048$) were identified with the forward selection procedure followed by the RDA analysis performed on the geographical data. Selected variables explained 39.03% of the variation observed in the EAB gut bacterial community (**Figure 2.4**). The three predictor matrices displayed 5.72% redundancy, whereas the shared contribution of phyllosphere bacteria and fungi showed the greatest effect on the gut bacterial community (22.79%). The six phyllosphere bacterial ASVs explained a modest 0.12% of the gut bacterial community structure, while the six leaf fungi ASVs explained 6.61% of the variation. The relationship between geographical site coordinates and the gut bacterial community was redundant, with the variation explained by the microbiome of the phyllosphere.

Tableau 2.2 **Details on the selected ASVs of the variance of bacterial et fungal communities associated to adult EAB guts. In both cases, the values correspond to the percentage of presence of the ASV taxa in all the samples. Only the ASV found to be indicators of a host are shown.**

ASV[1]	Phylum	Class	Genus	Guts (n=18)	Leaves (n=18)
Bacterial community					
ASVb869	Proteobacteria	Gammaproteobacteria	Unassigned	16,7%	33,33%
ASVb78	Actinobacteria	Actinobacteria	*Kineococcus*	11,1%	55,56%
ASVb299	Proteobacteria	Alphaproteobacteria	Unassigned	16,7%	33,33%
ASVb205	Proteobacteria	Alphaproteobacteria	*Methylobacterium*	27,8%	44,44%
ASVb105	Proteobacteria	Gammaproteobacteria	*Acinetobacter*	16,7%	44,44%
ASVb5	Actinobacteria	Actinobacteria	*Propionibacterium*	88,9%	100,00%
Asvf26	Ascomycota	Unassigned	Unassigned	5,6%	55,56%
Asvf145	Ascomycota	Unassigned	Unassigned	0,0%	44,44%
Asvf235	Ascomycota	Leotiomycetes	*Naevala*	5,6%	33,33%
Asvf90	Ascomycota	Sordariomycetes	*Diaporthe*	0,0%	38,89%
Asvf263	Ascomycota	Dothideomycetes	Unassigned	5,6%	33,33%
Asvf173	Ascomycota	Dothideomycetes	Unassigned	0,0%	38,89%
Fungal community					
Asvb52	Actinobacteria	Actinobacteria	Unassigned	27,8%	55,6%
Asvb633	Proteobacteria	Alphaproteobacteria	Unassigned	5,6%	50,0%
Asvb10	Proteobacteria	Gammaproteobacteria	*Escherichia*/*Shigella*	55,6%	44,4%
Asvb740	Proteobacteria	Betaproteobacteria	*Massilia*	0,0%	38,9%
Asvb896	Proteobacteria	Alphaproteobacteria	*Sphingomonas*	38,9%	66,7%
Asvb114	Actinobacteria	Actinobacteria	Unassigned	22,2%	50,0%
Asvb503	Proteobacteria	Alphaproteobacteria	*Methylobacterium*	27,8%	66,7%
Asvf486	Ascomycota	Dothideomycetes	*Phaeosphaeria*	5,6%	55,6%
Asvf292	Ascomycota	Dothideomycetes	Unassigned	0,0%	44,4%
Asvf8	Ascomycota	Dothideomycetes	*Phoma*	61,1%	100,0%
Asvf123	Ascomycota	Dothideomycetes	Unassigned	11,1%	55,6%

Seven bacterial ASVs: ASVb52 (Actinobacteria), ASVb633 (Alphaproteobacteria), ASVb10 (Gammaproteobacteria), ASVb740 (Betaproteobacteria), ASVb896 (Alphaproteobacteria), ASVb114 (Actinobacteria), and ASVb503 (Alphaproteobacteria), and four fungal ASVs: ASVf486, ASVf292, ASVf8, and ASVf123, all associated with the phyllosphere and belonging to the Dothideomycetes, explained the variation in the gut fungi. The forward selection procedure followed by the RDA analysis performed on the geographical data identified MEM14 ($F = 3.2058$, adj. $R^2 = 0.12$, $p = 0.001$), and MEM13 ($F = 2.11$, adj. $R^2 = 0.18$, $p = 0.031$) as significant variables explaining the variation in gut fungi.

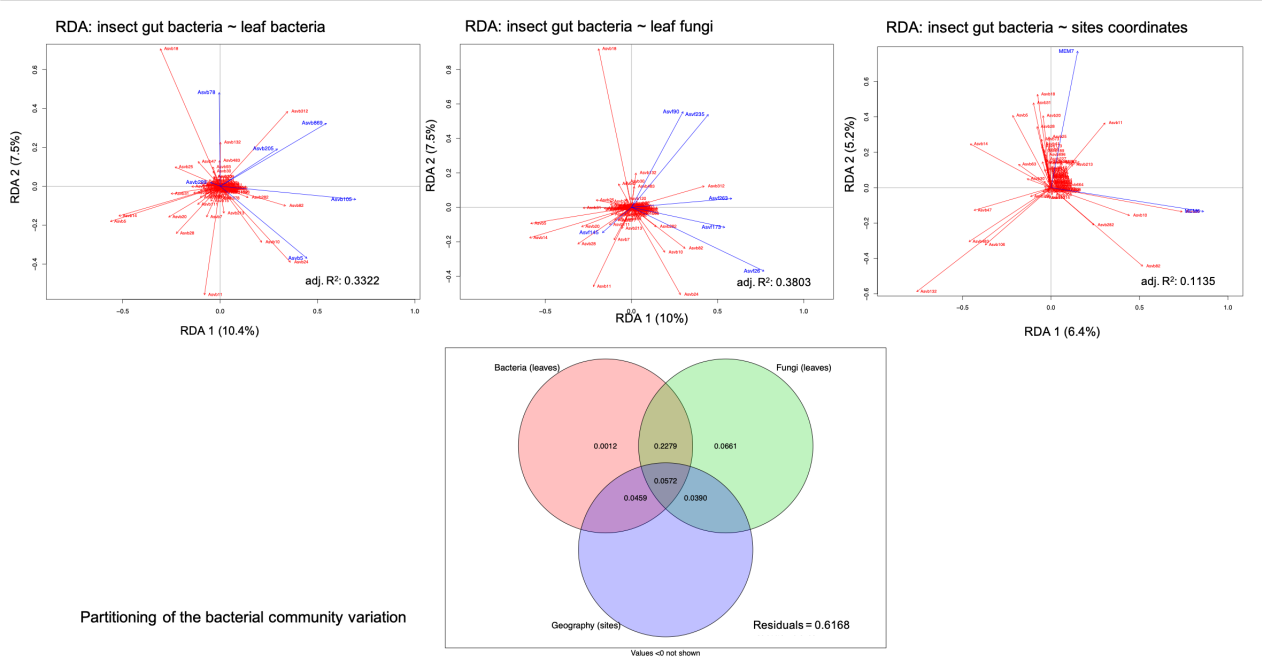


Figure 2.3 : Variation partitioning of the gut bacterial community among three predictor matrices: bacteria associated with leaves (Hellinger-transformed), fungi associated with leaves (Hellinger-transformed), and dbMEM spatial eigenfunctions generated from the geographic coordinates of the sampling sites (*Fraxinus* trees). The selected explanatory variables (leaves) are represented in blue, and the response variables (insects' gut) in red. Thus, the three explanatory matrices (bacteria, fungi, and geographic coordinates) are represented in individual RDA analyses. The adjusted R-square (adj.R^2) corresponds to the R^2 adjusted to the model containing all variables. The figure below the RDAs represents the partitioning variation analyses of the bacterial community associated with adult EAB gut.

The phyllosphere microbiome and the sampling site location explained 52.68% of the variation observed in the gut fungal community (**Figure 2.5**).

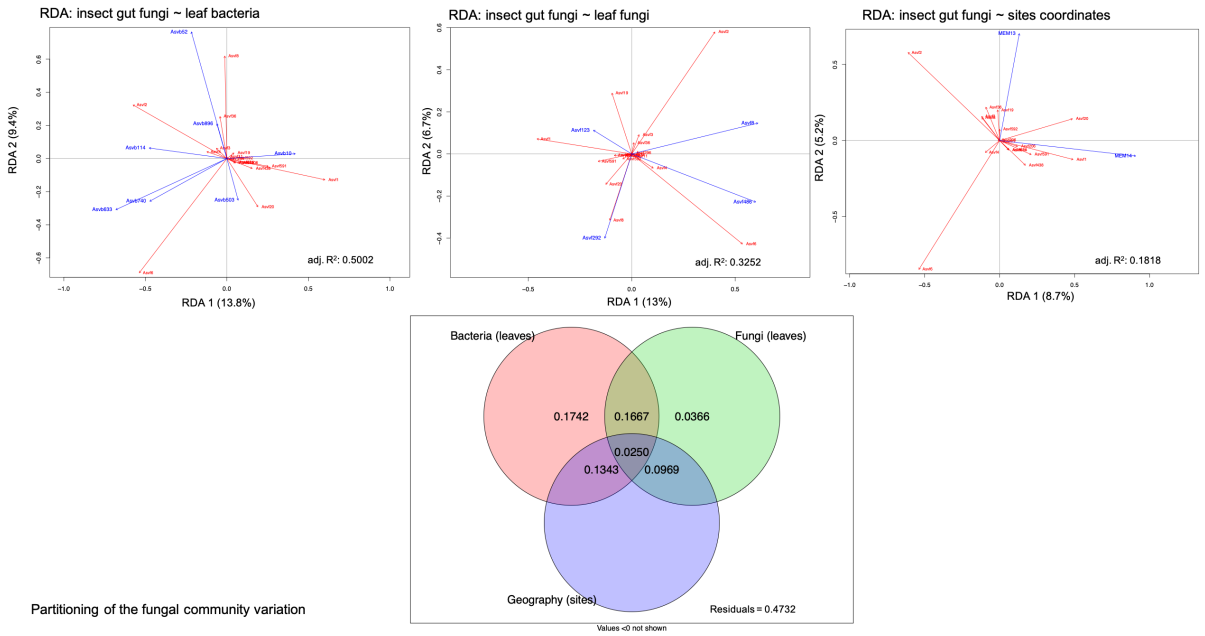


Figure 2.4 : Variation partitioning of the gut fungal community among three predictor matrices: bacteria associated with leaves (Hellinger-transformed), fungi associated with leaves (Hellinger-transformed), and dbMEM spatial eigenfunctions generated from the geographic coordinates of the sampling sites (*Fraxinus* trees). The selected explanatory variables (leaves) are represented in blue, and the response variables (insects' gut) in red. Thus, the three explanatory matrices (bacteria, fungi, and geographic coordinates) are represented in individual RDA analyses. The adjusted R-square (adj.R^2) corresponds to the R^2 adjusted to the model containing all variables. The figure below the RDAs represents the partitioning variation analyses of the fungal community associated with adult EAB gut.

The individual contribution of the phyllosphere bacterial community to the variation in the EAB gut microbiome was higher on the fungal community than on the bacterial community, with respectively 17.42% and 0.12%. The individual contribution of the geographic coordinates was not significant (<0%), while its common contribution to the bacteria and fungi was 13.43% and 9.69%, respectively. The common contribution of the three predictor matrices was 2.50%. The results indicate a stronger relationship between the ash phyllosphere microbiome and the gut fungi than the gut bacteria (**Figure 2.6** and **Table Suppl. 2.2**).

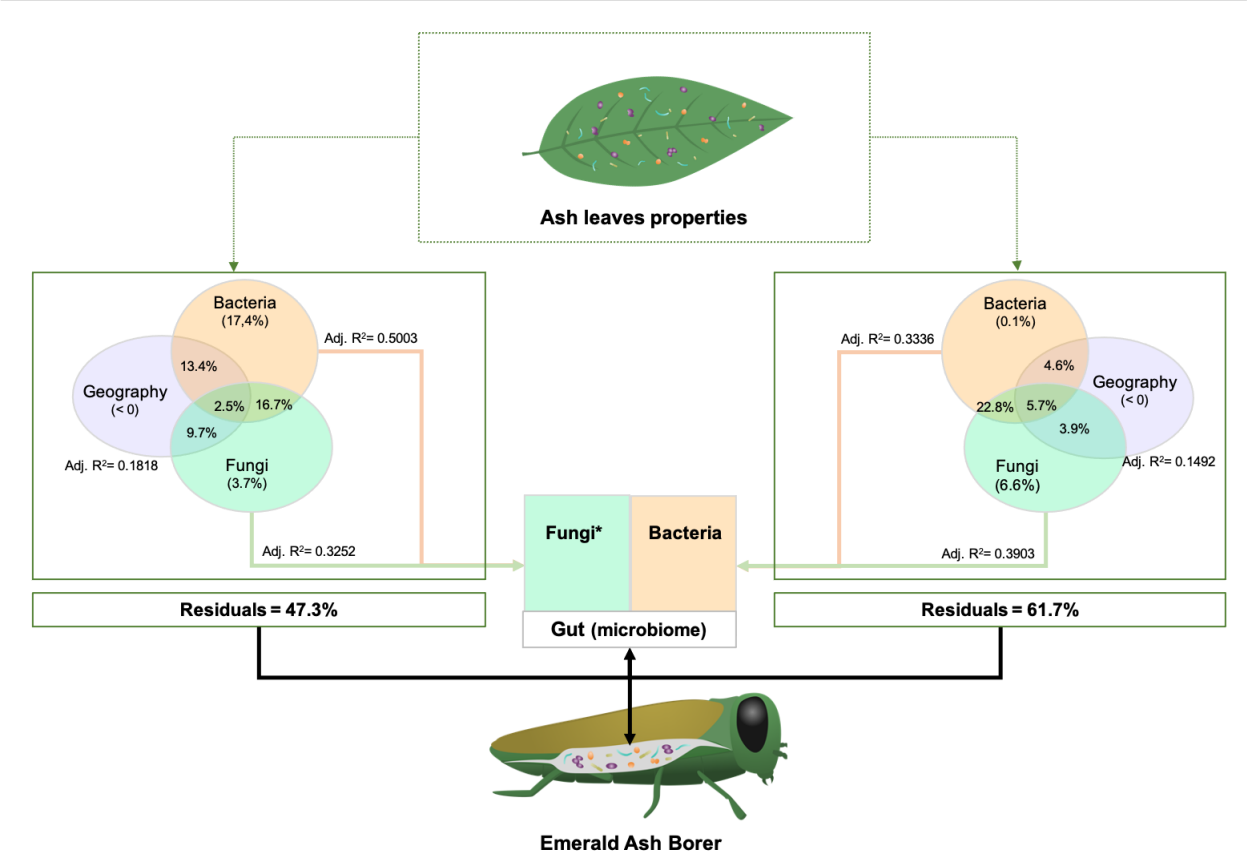


Figure 2.5 : Global representation showing variation partitioning analysis of microbial communities associated with leaves on the microbial communities associated with adult EAB gut. The host tree descriptors that have been found as explanatory matrices (bacteria, fungi, and geographic coordinates) are represented on the left portion (for the variation of the EAB gut fungal community), and on the right portion (for the variation of the EAB gut bacterial community). The dash box indicates the explanatory variable (properties of the host tree), and the values indicated in the circle correspond to the percentage of explanation. * After computing a forward selection (alpha < 0.05) prior to multiple regressions. The cellulose content significantly explained the species richness observed in the fungal community associated with the adult EAB gut

2.5 Discussion

The composition of the insect gut microbiome is influenced by diverse environmental drivers including developmental stage, diet, season, and plant defences. The role of diet in determining gut microbiome structure has been mostly related to the host's nutritional requirements for carbohydrates, vitamins, amino acids, and inorganic elements (Schowalter, 2016). These compounds are supplied either directly by leaves (carbohydrates and inorganic nutrients) or indirectly through microbial metabolism of leaf residues (amino acids, vitamins, and inorganic nutrients) (Schowalter, 2016). Microbiome studies are expected to shed light on the interplay

between the microbiome of the insect gut and host fitness, and therefore can increase basic knowledge of insect biology for the design of novel control methods. One limitation on this quest is the multiple filtering effects exerted on insect gut microbiome by abiotic and biotic influences. In the case of diet, it is reasonable to expect food to have both abiotic and biotic effects on the gut microbiome structure, but the impact of food-associated fungi and bacteria has not been documented. DNA sequencing approaches covering the overall structure of microbial communities, including living cells and relic DNA from dead cells or debris (Oliverio *et al.*, 2018) were proposed to bridge this knowledge gap. Although the distribution and relative abundance of living cells has been directly linked to host fitness, relic DNA is a legacy from the environment, and Lennon *et al.* (2018) maintained that relic DNA might be one of the largest pools of nucleic acids in the biosphere. In this study, disentangling the contribution of such a legacy on the host microbiome structure was made possible by selecting EAB adults known to be *Fraxinus spp.* host specialists. Both insect and host tree were examined to evaluate the covariation between host tree biotic and abiotic features and insect gut microbiome structure.

Dissimilarities between ash leaf and EAB gut biotopes were reflected in variations in the bacterial and fungal communities and indicator ASV levels (Fig. 3 and Table 1). The phyllosphere is exposed to solar radiation, pollution, and insect herbivory (Humphrey & Whiteman, 2020; Kadivar & Stapleton, 2003; Kembel *et al.*, 2014; Thapa & Prasanna, 2018; Vorholt, 2012; Yutthammo *et al.*, 2010), while most insect gut environments are alkaline and protected from environmental influences except through feeding. In spite of these contrasting conditions, the microbiomes of both biotopes exhibited similar diversity according to the computed diversity indices. However, the significantly lower phylogenetic diversity in the gut fungal community, brought up the question whether the phylogenetic diversity could be related to the ecological process happening in related the fungal community of each biotope. Fungi and insects have been largely documented to have mutualistic relationships, such as in those insect species that feed on wood or woody detritus. The fungal activities that benefit insects include the breakdown of cellulose or other recalcitrant molecules and protection through synthesis of repellents and antimicrobial metabolites (Biedermann & Vega, 2020).

Interestingly, the ash leaf cellulose content was related to the fungal community associated with the adult EAB gut. Although some insects can synthesize their own cellulases, the enzymatic activity of symbiotic microorganisms is often needed for complete cellulolysis (Fischer *et al.*, 2013; Schowalter, 2016). Some microorganisms, including fungi, have been shown to possess enzymes capable of hydrolysing a wide variety of plant materials (Watanabe & Tokuda, 2010). Mittapalli

et al. (2010) even found microbial transcripts coding for cell-wall degrading enzymes in the transcriptome of the EAB larvae midgut. The variation partitioning approach highlighted the covariations of bacterial and fungal profiles associated with both biotopes.

Compelling evidence supports the idea that environmental microorganisms play a major role in shaping the insect gut microbiome. For instance, gut colonization of a lepidopteran species (*Pieris brassicae*) by the bacterium, *Pseudomonas protegens*, commonly associated with plants and soil is promoted by the disruption of the commensal microbiome in the insect gut induced by the bacterium (Vacheron *et al.*, 2019). The potential effects of food-associated microbes on gut microbiome has been documented in some fish species, confirming the role of food associated microbes in the prediction of population differences in the gut microbiota (Smith *et al.*, 2015). To the best of our knowledge, there has been no previous report on the effects of phyllosphere microbiome on the gut microbial community of a herbivorous insect. Here, the fungal and bacterial communities associated with the ash phyllosphere explained 53 and 48% of the variation observed in the adult EAB gut microbiome. We found that the gut bacterial community might be less sensitive to the phyllosphere microbiome than the gut fungal community. A previous work by Zhang *et al.* (2018) investigating the bacterial and fungal communities in the gut of adult *Agrilus mali* ingesting different diets, revealed a greater change in the fungal community than in the bacterial. Based on our results, we would suggest that the changes in the fungal community associated with the *Agrilus mali* gut could be due to food-associated microorganisms. Both environmental fungi as a complement of insect diet and dissemination of gut-associated fungi in the environment can explain the covariation between the EAB gut and the phyllosphere microbiomes. However, more investigation is needed to better understand the potential relationships between insect gut fungi and the fungi associated with the phyllosphere of food plants, as well as the influence of gut bacteria on insect metabolism. Variation partitioning analysis showed significant covariation of the bacterial and fungal communities associated with the ash phyllosphere and the geographic coordinates of the host sites with the EAB gut microbiome. In each variation partitioning analysis, the joint redundancy of the phyllosphere bacterial and fungal communities was higher than the unique proportions explained by bacterial and fungal communities (**Figure 2.6**). High redundancy suggests biological interactions between bacteria and fungi occurring in both leaf and gut biotopes (Agler *et al.*, 2016). Greater knowledge about the mechanism of microbial interactions may improve our understanding of the biological interactions among coexisting microbial communities.

The successful adaptation of insects to environmental changes has been reported in many papers. According to Gupta and Nair (2020), microbes may have played a crucial role in insect survival, and gut bacteria may facilitate faster adaptations of an insect host to a changing environment. Schowalter (2016) also reported that insect responses to environmental conditions determined their survival, reproduction, and fitness. Variation in the microbial composition in the EAB gut is certainly related to their adaptation to the host tree properties. As to our results demonstrating the covariation of the phyllosphere-host microbiomes, future investigations are needed to elucidate the mechanism behind the microbiome variation and, more importantly, the consequences of that variability on the adaptation of insects facing environmental changes. Understanding how beneficial functional traits can be provided to EAB through ingestion of microbiota associated with ash leaves could provide valuable insights into the contribution of environmental microorganisms to insect fitness, and how this might be manipulated to human advantage in the case of pest species such as the EAB.

2.6 Materials and methods

2.6.1 Sampling site locations, insect collection, and leaf sampling

Site locations. Eighteen white ash trees, *Fraxinus americana* L., separated by a radius of at least 200 m from any other ash tree, were randomly selected in the Lanaudière region North-East of Montréal, QC, Canada (**Figure 2.1**). White ash trees were identified by CG using a tree identification key (AFSQ, 2018).

Sampling of adult EABs. Adult EABs were collected using 12-funnel green Lindgren traps (Synergy Semiochemicals Corp., Burnaby, BC, Canada) deployed in the highest third of the canopy of each selected ash tree using a catapult (North Big Shot, Sherriltree, Greensboro, NC, USA) to set up a permanent halyard system for regular trap monitoring. Based on previous observations on adult flight activity and the calculation method of degree-days (Mogouong *et al.*, 2020), traps were deployed and monitored after a period of 341.7 degree-days (base 10°C) was reached. Adults were collected twice weekly in 2018, from June 21st to August 3rd. For each selected tree and each sampling date, insects were transferred into a disinfected plastic container using sterilized tweezers and supplied with fresh ash tree leaves as substrate and food source. Containers were kept on ice during transport to the laboratory. Each healthy insect was immediately transferred into labelled sterile 1.5 mL microfuge tubes, frozen at -80°C, and stored until DNA extraction. The number of insects collected per tree ranged from two to 100 EAB adults.

Leaf sampling. For each selected tree, for each selected tree, several leaves surrounding the trap were collected using a 10-m aluminium extendable pole tree pruner. Separate samples were placed in sterile bags and transported on ice to the laboratory at the fourth sampling week. Subsamples of each tree corresponding to five leaves randomly selected were then stored at -80°C before DNA and metabolite extractions. Until digitalization, foliar voucher samples are available to Guertin's laboratory for further consultation.

Several site locations were visited to collect leaf samples and deployed traps to catch insects. For each place where plants and EAB adults have been collected, we obtain the authorization of private landowners and municipalities before collecting and using them in this study.

Our study complies with relevant institutional, national, and international guidelines and legislation.

2.6.2 DNA extraction

For each tree, insect DNA extraction was performed on two randomly selected EAB adults. The wings were first removed with sterile tweezers and scissors, and the exoskeleton was sterilized by agitating (Fisher Vortex Genie 2, Ottawa, ON, Canada) the beetle in 1 mL of 70% ethanol for 1 min. Beetles were rinsed with 1 mL of sterile water by vortexing for 30 s. Dissection was performed in sterile phosphate-buffered saline, the two guts were pooled, and total genomic DNA was extracted by the mechanical lysis as described by Mogouong *et al.* (2020). For each leaf sample, endophytes and epiphytes were processed together by grinding five leaflets corresponding to the apical leaflet of five leaves randomly selected (**Figure Suppl. 2.4**) in liquid nitrogen and using the MoBio PowerSoil DNA isolation kit (Qiagen, Toronto, ON, Canada). DNA extracts from insect guts and leaves were purified using a PowerClean Pro DNA clean-up kit (Qiagen, Venlo, the Netherlands). DNA concentration was estimated using the Quant-iT PicoGreen dsDNA assay kit (Invitrogen, Life Technologies, Burlington, ON, Canada) following the manufacturer's instructions.

2.6.3 DNA amplification by PCR, amplicon sequencing and data processing

The v6-v8 region of the 16S rRNA gene of bacteria and the ITS2 region for fungi were sequenced using, respectively, the primers B969F-CS1 (5'-ACGCGHNRAACCTTACC-3') and BA1406R-CS2 (5'- ACGGGCRGTGWGTRCAA-3') (Comeau *et al.*, 2011) and the primers ITS3_KYO2 (5'-GATGAAGAACGYAGYRAA-3') and ITS4_KYO3 (5'-CTBTTVCCKCTTCACTCG-3') (Toju *et al.*, 2012) for both insect and leaf DNAs. Sequencing was done using an Illumina Miseq with 2 x 250

bp paired-ends at the Quebec Genome Innovation Centre (Sainte-Justine Hospital, Montreal, QC, Canada). Sequence data were processed with USEARCH, v10.0.240 (Edgar, 2010b; Edgar, 2016), using a pipeline for constructing amplicon sequence variants (ASVs) (Callahan *et al.*, 2017; Glassman & Martiny, 2018). The ASV matrix returned frequencies of ASVs per sample. The main steps of the pipeline included quality filtering, dereplication by finding unique sequences, merging of paired-end reads (total length of 438-484 bp, for bacteria, or 379-450 bp, for fungi), clustering of ASVs, creation of the ASVs abundance table (an ASV was generated with at least eight sequences with 100% identity) after removing the chimeras, and lastly taxonomy assignment. The taxonomy assignment was completed using the RDP (Ribosomal Database Project) (Cole *et al.*, 2014; Wang *et al.*, 2007) v.16 training set of 16S rRNA genes for bacteria and the RDP Warcup training set v.2 for fungi. Only assignments >80% were considered. The non-assigned ASVs were identified by the label, 'non-assigned'. The bacterial ASV table comprised 1073 ASVs represented by 326,583 sequences, while the fungal ASV table comprised 602 ASVs clustered in 1,431,817 sequences.

2.6.4 Extraction of leaf phytochemicals

Total soluble proteins. Extraction of total soluble protein was performed as described by Chen and Poland (2009) and Bi *et al.* (2003). Briefly, 40 mg of ash leaves were ground in liquid nitrogen and transferred to tubes containing 5 mL of 0.1 M ice-cold phosphate buffer (pH 7.0) containing 1% polyvinylpolypyrrolidone. After vortexing and centrifugation at 10,000 x g for 10 min at -2 °C, the total protein content in the supernatants was determined using the bicinchoninic acid (BCA) protein assay. The absorbances of the reaction mixture were read at 562 nm using a Tecan Infinite M1000 Pro microplate reader (Tecan US, Morrisville, NC, USA). The protein concentrations were calculated from a standard curve of bovine serum albumin (BSA) and expressed as µg of protein per g of dry leaf mass (µg/g DW) after subtracting the water content of each sample.

Total phenolics. Total phenolic content was determined as described by Torti *et al.* (1995) and Hagerman (1988). After removing prominent veins, fresh leaves were ground, and 200 mg was extracted in 5 mL of 70% acetone. The samples were sonicated for 30 min at 4°C and then centrifuged at 16,000 x g for 15 min at 4°C. The amount of extract corresponding to the pooling of three successive extractions was used for the analysis. The total phenolic content was determined using the Folin-Ciocalteu reagent assay described by Beauchemin *et al.* (2012) with some modifications. First, 50 µL of each extract was mixed with 2.5 mL of Folin-Ciocalteu reagent and incubated in a 40 °C water bath for 8 min. Then, 1 mL of sodium carbonate 1M was added

and samples were incubated at 40°C for one hour in the dark. Aliquots of each sample's reaction products were transferred to a 96-well microplate and absorbance was measured at 765 nm using a SpectroStar Nano spectrophotometer (BMG Labtech, Germany) with gallic acid standards to determine the total phenolic content expressed in mg of gallic acid equivalents per g of dry leaf mass, based on the water content of each sample (mg/g DW).

Chlorophyll and carotenoids. Using the dimethyl sulfoxide (DMSO) method described by Garg (2012), the chlorophyll and carotenoid content were determined. Briefly, 50 mg of fresh leaf tissue, cut in pieces of approximatively 1 x 1 cm, were placed in a tube containing 7 mL DMSO, and incubated at 65°C for 3 h. After removing the leaves, the extracts were transferred to graduated tubes, and the volume was adjusted to 10 mL with DMSO. Absorbance was measured at 480 (A_{480}), 645 (A_{645}), and 665nm (A_{665}) using a HACH DR2800 spectrophotometer (Hach Canada, London, ON, Canada), and chlorophyll and carotenoid concentrations were calculated based on the following formulas from Wellburn (1994):

Chlorophyll a ($\mu g\ mL^{-1}$) = $12.19*A_{665} - 3.45*A_{645}$

Chlorophyll b ($\mu g\ mL^{-1}$) = $21.99*A_{645} - 5.32*A_{665}$

Carotenoids ($\mu g\ mL^{-1}$) = $(1000*A_{480} - 2.86*\text{chlorophyll a} - 129.9*\text{chlorophyll b})/221$

The final concentrations were expressed as mmol per g of dry mass (mmol/g DM).

Soluble carbohydrates and starch. Soluble carbohydrates and starch contents of leaves were determined as described by Marquis *et al.* (1997). For each sample, fresh leaves were ground in liquid N_2 and 30 mg was placed in a centrifuge tube containing 1.5 mL ethanol 80% (v/v) and incubated in a water-bath at 80°C for 8 min. Following centrifugation (10,000 x g at 2°C) for 10 min, the supernatants were collected (extract A), and the extraction procedure repeated one time (extract B). The soluble carbohydrates were pooled (extracts A and B) for quantification. The pellets containing insoluble starch were dehydrated overnight at 50°C. The tube containing the pellet was rinsed with 80% ethanol to remove any soluble carbohydrates. The starch was hydrolysed by addition of 2.5 mL of acetate buffer (0.2 M, pH 4.5) to the pellets and incubation in boiling water for one hour. After cooling to room temperature, 2 mL of acetate buffer and 1 mL of glucoamylase (20 units/mg at pH 6.0) at 0.5% (Bio Basic Inc., ON, Canada) were added to the mixture, incubated at 55°C for 8 h, and supernatants were collected (extract C). The soluble carbohydrate extracts (A and B pooled) and the starch hydrolysate (extract C) were then filtered through 0.45 µm syringe filters (Bio Basic Inc. ON, Canada). Soluble sugars were quantified and analysed using a high-performance liquid chromatograph (HPLC, Shimadzu) equipped with a

pump (LC-10VP), an autosampler (SIL-10AXL), a column oven (CTO-10ASVP) at 65 °C, a refractive index detector (RID-10A), a system controller (SCL-10AVP), and an Aminex HPX-87H column (300 mm, 7.8 mm). The mobile phase consisted of 5 mM H_2SO_4 at 0.6 ml min^{-1}. The soluble carbohydrate extract concentration was determined by quantifying specific sugars (sucrose, glucose, fructose) against their respective standards (D-sucrose, D-glucose, D-fructose). Starch concentration was measured by quantifying hydrolysed products (glucose) against the glucose standard and expressed in glucose equivalents. Finally, the non-structural carbohydrate (TNC) content was computed by combining the content of the three extracted sugars (glucose, fructose, and sucrose) and the hydrolysed starch (D-glucose equivalents). The sugar fraction was expressed in mg per g of dry mass (mg/g DM).

C and N content of leaves. Leaves were dried at 70°C for 72 h and ground. The carbon and nitrogen content, and the carbon-to-nitrogen ratio were determined by combustion using an Elementar Vario Micro Cube (Elementar, Germany) and acetanilide as the standard. The amount of carbon and nitrogen were expressed as percentages.

Leaf degradability. The degradability of leaf fibre was estimated by extracting tissue fractions as described by García *et al.* (2003) and using an ANKOM fibre analyzer (Ankon Technology, Macedon, NY, USA). The fractions extracted, including neutral detergent fibre (NDF), acid detergent fibre (ADF), acid detergent lignin (ADL), cellulose and hemicellulose, were expressed as percent of tissue (García *et al.*, 2003; Van Soest *et al.*, 1991).

2.6.5 Statistical analyses

Richness and diversity. The raw data was used to assess the α-diversity expressed as species richness and ecological diversity indices. The species richness and diversity were investigated by computing the species richness and two diversity indices (Shannon and Simpson, with corresponding Hill numbers) using the diversity() function available in the 'vegan version 2.5-6' package (Oksanen *et al.*, 2019) and following the procedure proposed by Borcard *et al.* (2018). Faith's phylogenetic diversity (PD) calculation allowed us to determine the environmental diversity patterns (Kembel *et al.*, 2011). The PD, corresponding to the sum of the branch lengths belonging to the minimum spanning path, was computed based on cladistic information (Faith, 1992) using the PD() function available in the 'picante version 1.8.1' package (Kembel *et al.*, 2010). The computed diversity indices of microbial communities associated with the insect guts and the leaves were compared using Wilcoxon tests.

Microbiome composition variation. Among sites, the ß-diversity defined as the variation in community composition among sampling sites (i.e., the 18 sampled trees) was investigated for the microbial communities in the adult EAB guts' and on the leaves, without reference to an explicit gradient. The ß-diversity was determined by three complementary approaches after removing ASVs whose frequency in the ASV table was < 0.005% and performing the Hellinger transformation. The first approach, proposed by Legendre and De Cáceres (2013), calculates the local contributions of each site to β-diversity (LCBD), defined as "comparative indicators of the uniqueness of a site" and based on the ASVs' composition and the species contributions to the β-diversity (SCBD), which represents the level of contribution of each ASV to the overall β-diversity. LCBD and SCBD were computed using Hellinger-transformed ASV frequencies with the beta.div() function available in the 'adespatial version 0.3-8' package (Dray *et al.*, 2020). The final values of SCBD and LCBD, ranging from 0 to 1, indicated the importance of the species and the sites in the overall β-diversity for each microbial community. A prerequisite of the method is the evaluation of the total variance of the Hellinger-transformed ASV matrix, called BD_{Total}, which represents an estimate of the ß-diversity.

$BD_{Total} = SS_{Total}/(n-1)$, where SS_{Total} corresponds to the total sum of squares for Hellinger-transformed ASV abundances, and n is the total number of sites. BD_{Total} was then decomposed into LCBD (relative contribution of the site i to the ß-diversity) and SCBD (relative contribution of ASV column j to the ß-diversity) indices computed as follows:

$LCBD = SS_i/SS_{Total}$, where SS_i corresponds to the contribution of the site i to the overall ß-diversity.

$SCBD = SS_j/SS_{Total}$ where SS_j corresponds to the contribution of ASV column j to the overall ß-diversity.

The second approach examines the distance between the two biotopes (adult EAB guts and ash leaves) based on the ASV frequencies of the microbial members of the two groups. We performed a distance-based redundancy analysis (db-RDA with Bray-Curtis distance) by using the capscale() function of the 'vegan' version 2.5-6 package (Oksanen *et al.*, 2019). Briefly, a Bray-Curtis distance matrix of samples from adult EAB guts and the phyllosphere was constructed and used to determine the principal coordinates that were then used to analyse its relationship with each biotope using RDA. The model's significance was validated by a permutation test using the function anova.cca() of the 'vegan' version 2.5-6 package. The third ß-diversity analysis was based on the concept of indicator species proposed by De Cáceres *et al.* (2010) identifying ASV indicators that could reflect the state of the environment using the multipatt() function available in the 'indicspecies' version 1.7.9 package (De Cáceres & Jansen, 2020).

Relationships between ash leaf phytochemicals and the microbiome in the adult EAB gut. Relationships between the extracted phytochemicals and the diversity indices computed for the microbial communities associated with the adult EAB gut were analysed using the permutational selection approach proposed by Blanchet *et al.* (2008) based on the forward selection of explanatory variables in a linear regression function available in the 'vegan' package. Briefly, the forward.sel() function of the 'adespatial' package selects the best explanatory variable(s) in an explanatory matrix (phytochemical dataset) which suits a predefined model (ASV frequencies) with a preselected significant p value (0.05). Selected molecules were used in a regression model and tested with a permutation test using the function anova.cca() available in the 'vegan' package. The relationship between the computed LCBD indices and the phytochemicals were investigated using the same approach.

Partitioning of the variation. Relationships between the microbial communities related to the adult EAB guts and their food source, represented by ash leaf phytochemicals and associated microbial communities, were analysed following the approach proposed by Borcard *et al.* (1992) and Peres-Neto *et al.* (2006). Briefly, data matrices of predictors related to the ash leaves (bacterial community, fungal community, phytochemicals, and geographic positions of sample sites (18 sampled ash trees) were tested as predictor variables of the bacterial and fungal communities in EAB guts. The ASV matrices utilized for variation partitioning contained a subset of ASVs whose occurrence was observed in at least 30% of the samples. This arbitrary cutoff was selected to eliminate rare ASVs with sparse distribution profiles. Thus, each predictor matrix was used in a partial RDA after a preliminary Hellinger transformation of the ASV abundance matrices. The sampling sites' geographic coordinates were included as a potential predictor matrix after transforming them into distance-based Moran's eigenvector maps (dbMEMs). The forward.sel() function was used to select the significant dbMEMs before the test of significance in a partial RDA and the variance partitioning analysis. Lastly, the variation partitioning of the microbial communities in the adult EAB gut enabled us to disentangle the contribution of each evaluated matrix. The variation partitioning was separately performed for the bacterial and fungal communities associated with the adult EAB gut by using the varpart() function of the 'vegan' package and following the procedure of Borcard *et al.* (2018) (**Figure 2.7**).

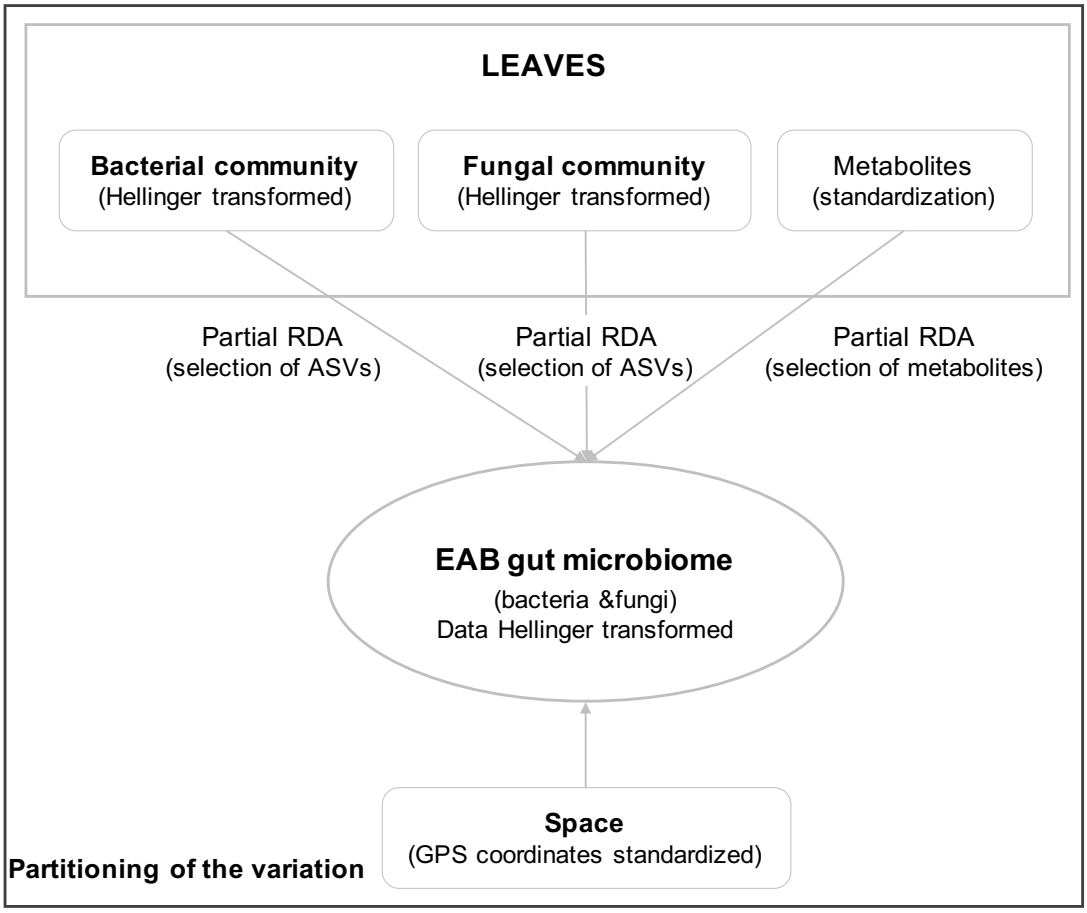


Figure 2.6 : Step-by-step representation of the two 'variance partitioning' analyses, including the variables datasets concerned. The circles show the response dataset, and each rectangle corresponds to a dataset of predictor variables. The response variables were transformed before the variance partitioning analysis.

Data Availability

Raw data is available in the Sequence Read Archive of the National Centre for Biotechnology Information (SRA, NCBI) under the BioProject number (PRJNA645095) and the SRA (SRP271139)

Acknowledgements

We thank Narin Srei (INRS-Centre Armand-Frappier Santé Biotechnologie) and Adrien Frémont (Université de Montréal, Montréal, Canada) for providing their technical assistance respectively for the field sampling across the large area covered by this study and for the extraction and quantification of the leaves' phytochemicals. This work was supported by INRS-NRCan-GDG Research Project to C.G. J.M. was supported by a Wladimir A. Smirnoff Fellowship and INRS Ph.D. Fellowship.

2.7 Supplementary materials

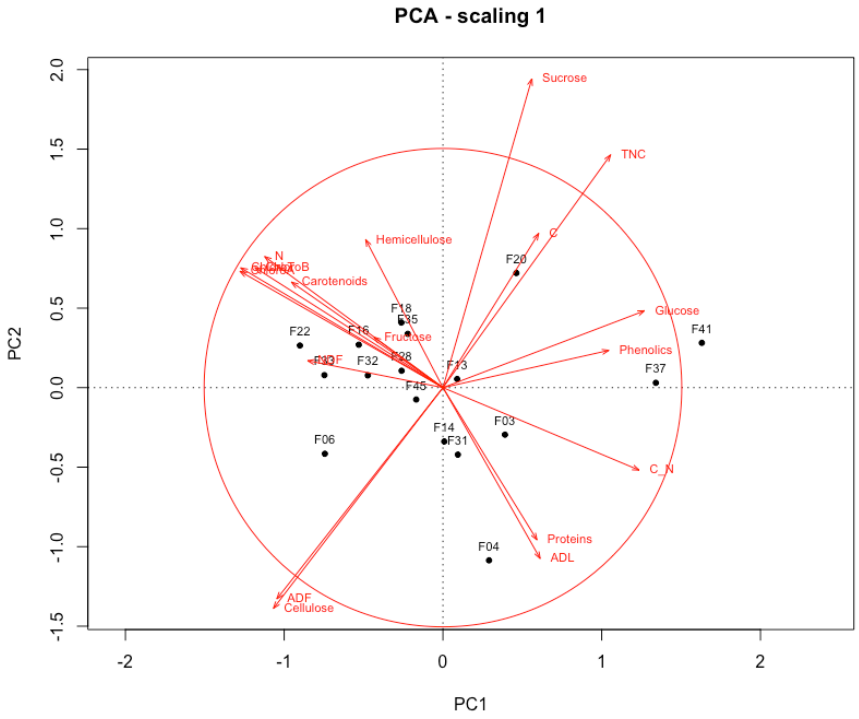
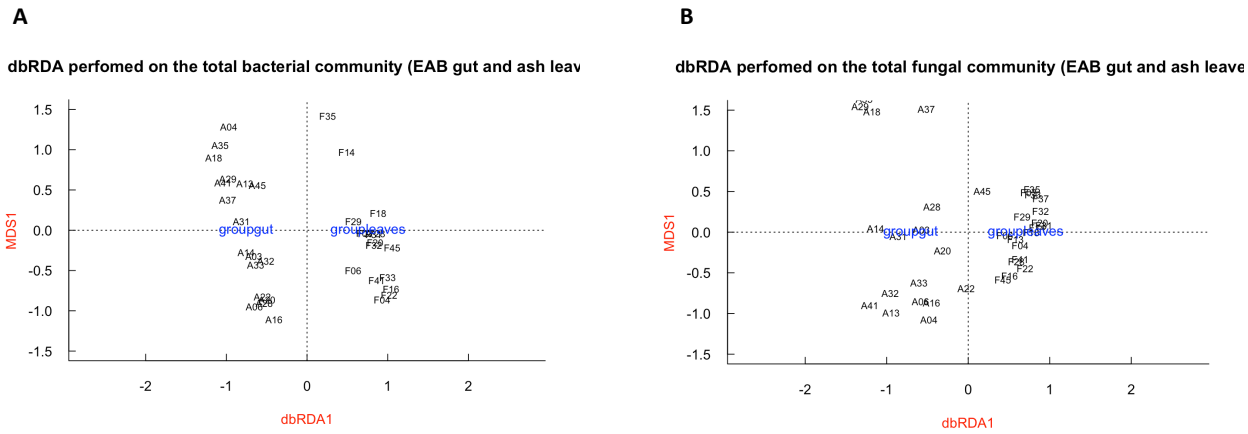


Figure Suppl. 2.1: The principal component analysis (PCA) performed with the eighteen phytochemicals extracted from ash leaves. The values were standardized and scaled before prior to the PCA analysis. Four molecules appeared as contributing the most to the dispersion of the sites in reduced space (two-dimensions): the cellulose, the acid detergent fibre fraction (ADF), the sucrose, and the total non-structural carbohydrates (TNC).



Bacteria

	Df	SumOfSqs	F	p(>F)
Model	1	1.3375	5.7853	0.001***
Residual	34	7.8605		

Fungi

	Df	SumOfSqs	F	p(>F)
Model	1	1.3375	5.7903	0.001***
Residual	34	7.8536		

Figure Suppl. 2.2: db-RDA analysis of bacterial (A) and fungal (B) communities associated to adult EAB gut and to leaves.

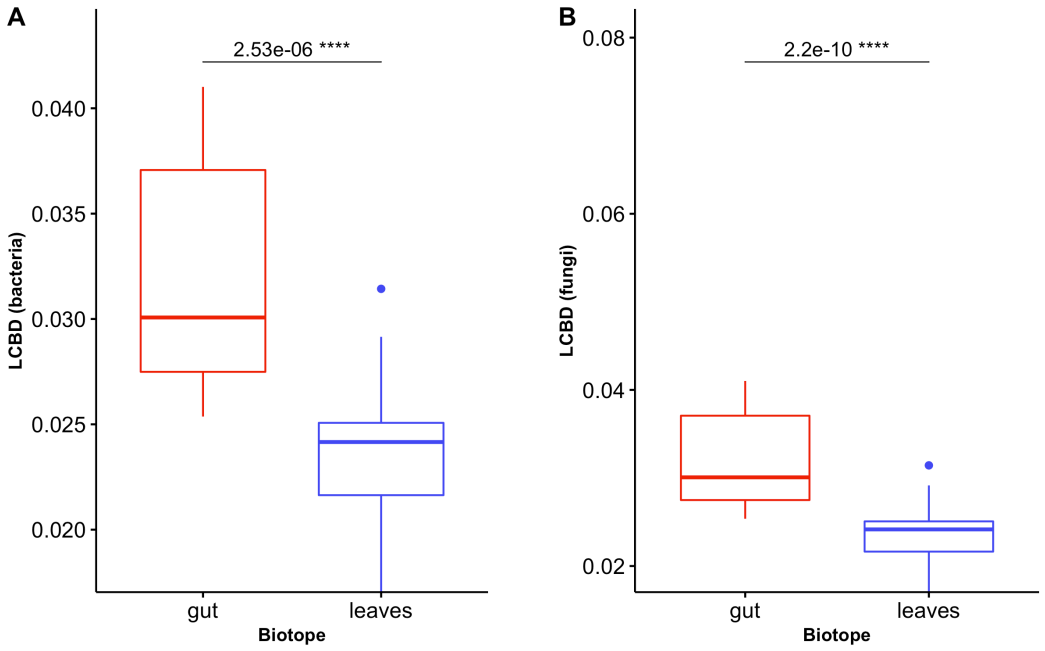


Figure Suppl. 2.3: Comparison of the LCBD values between the EAB and the leaf microbial communities showing a significant difference between the two communities.

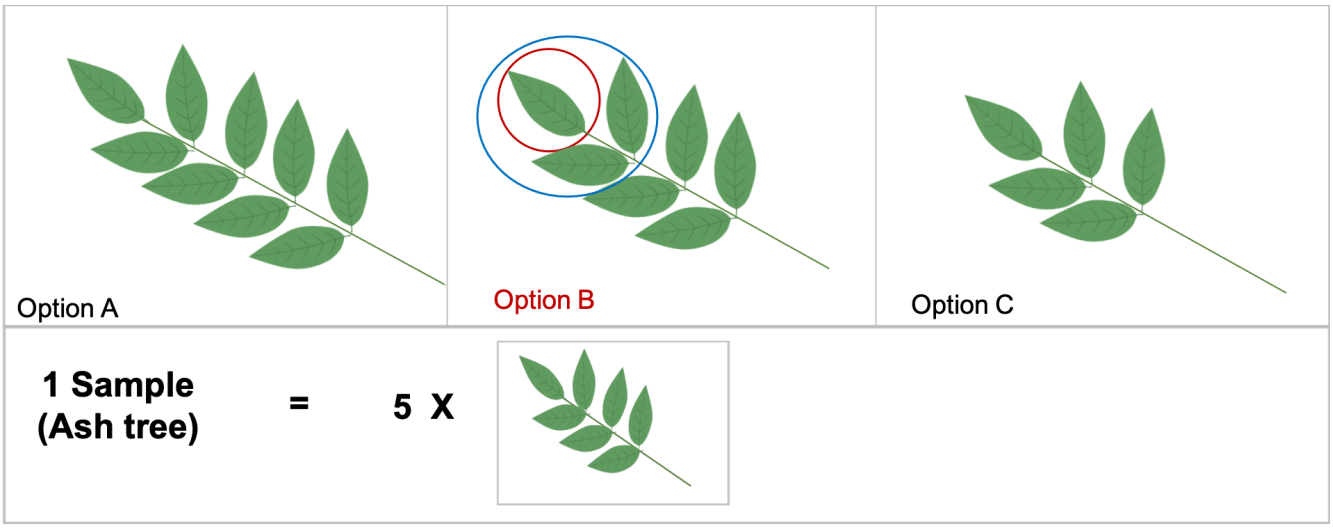


Figure Suppl. 2.4: For this study *Fraxinus americana* was used. Each leaf has five, seven or nine leaflets. For each leaf sample five apical leaflets of five randomly selected leaves were pooled, then used for DNA extraction (in red circle). For each sample the leaflets in the blue circle were used for phytochemicals extraction. During the collection of leaf samples, we obtained three different leaf structures, and we only considered the option B for the sampling.

Table Suppl. 2.1: The thirty-three bacterial ASVs assigned at the genus level using Ribosomal Data Base with a confidence level cutoff at 80% identified only in one habitat.

Only found in the adult EAB gut				
Genus	Phylum	Class	Names	Counts
Ewingella	Proteobacteria	Gammaproteobacteria	Asvb49	2462
Cryptococcus	Basidiomycota	Tremellomycetes	Asvf116	1961
Taibaiella	Bacteroidetes	Sphingobacteria	Asvb82	1650
Cordyceps	Ascomycota	Sordariomycetes	Asvf97, Asvf262, Asvf270	1487
Luteibacter	Proteobacteria	Gammaproteobacteria	Asvb63	1434
Coprinellus	Basidiomycota	Agaricomycetes	Asvf158	1036
Pseudochrobactrum	Proteobacteria	Alphaproteobacteria	Asvb129	813
Anoxybacillus	Firmicutes	Bacilli	Asvb233	670
Knoellia	Actinobacteria	Actinobacteria	Asvb283, Asvb494	658
Shewanella	Proteobacteria	Gammaproteobacteria	Asvb224	567
Exiguobacterium	Firmicutes	Bacilli	Asvb407, Asvb561	528
Herbiconiux	Actinobacteria	Actinobacteria	Asvb265	511
Lewia	Ascomycota	Dothideomycetes	Asvf197	499
Gaiella	Actinobacteria	Actinobacteria	Asvb242	498
Rubrobacter	Actinobacteria	Actinobacteria	Asvb324	424
Clostridium_sensu_stricto	Firmicutes	Clostridia	Asvb575, Asvb522	334
Iamia	Actinobacteria	Actinobacteria	Asvb298	332
Pluralibacter	Proteobacteria	Gammaproteobacteria	Asvb209	324
Aquitalea	Proteobacteria	Betaproteobacteria	Asvb347	313
Pelomonas	Proteobacteria	Betaproteobacteria	Asvb270	303
Annulohypoxylon	Ascomycota	Sordariomycetes	Asvf257	298
Methylocaldum	Proteobacteria	Gammaproteobacteria	Asvb460	276
Mesorhizobium	Proteobacteria	Alphaproteobacteria	Asvb498, Asvb674	271
Gemmiger	Firmicutes	Clostridia	Asvb413	225
Globicatella	Firmicutes	Bacilli	Asvb470	215
Morganella	Proteobacteria	Gammaproteobacteria	Asvb430	198
Inocybe	Basidiomycota	Agaricomycetes	Asvf333	184
Leptotrichia	Fusobacteria	Fusobacteria	Asvb532	179
Elizabethkingia	Bacteroidetes	Flavobacteria	Asvb523	173
Arthrobacter	Actinobacteria	Actinobacteria	Asvb459	172
Okibacterium	Actinobacteria	Actinobacteria	Asvb604	172
Actinomyces	Actinobacteria	Actinobacteria	Asvb569	166
Prevotella	Bacteroidetes	Bacteroidia	Asvb529	162
Phanerochaete	Basidiomycota	Agaricomycetes	Asvf407	136
Psathyrella	Basidiomycota	Agaricomycetes	Asvf420	124
Larkinella	Bacteroidetes	Cytophagia	Asvb611	115
Pseudoxanthomonas	Proteobacteria	Gammaproteobacteria	Asvb679	101
Only found in the leaves				
Diaporthe	Ascomycota	Sordariomycetes	Asvf90, Asvf326	1530
Sclerostagonospora	Ascomycota	Dothideomycetes	Asvf168	452
Coniothyrium	Ascomycota	Dothideomycetes	Asvf209	421
Diatrypella	Ascomycota	Sordariomycetes	Asvf352	230
Preussia	Ascomycota	Dothideomycetes	Asvf275	199
Neosetophoma	Ascomycota	Dothideomycetes	Asvf316	185
Drechslera	Ascomycota	Dothideomycetes	Asvf322	158
Ochrocladosporium	Ascomycota	Dothideomycetes	Asvf394	128
Lophodermium	Ascomycota	Leotiomycetes	Asvf410	104

[1] Asvb, ASV assigned to bacteria; Asvf, ASV related to fungi.

Table Suppl. 2.2: Summary extracted from the three partial RDA results

Bacterial community associated with adult EAB guts				
	Selected ASVs	Adj. R^2	F	*p*-value
Bacteria (leaves)	Asvb105 Asvb869 Asvb78 Asvb299 Asvb205 Asvb5	0.3336	2.335	0.001
Fungi (leaves)	Asvf26 Asvf145 Asvf235 Asvf90 Asvf263 Asvf173	0.3903	2.707	0.001
Phytochemicals	--	--	--	--
Geographical position	MEM6 MEM7	0.1135	2.024	0.002

Fungal community associated with EAB guts				
	Selected ASVs	Adj. R^2	F	*p*-value
Bacteria (leaves)	Asvb52 Asvb633 Asvb10 Asvb740 Asvb896 Asvb114 Asvb503	0.501	3.289	0.001
Fungi (leaves)	Asvf486 Asvf292 Asvf8 Asvf123	0.3252	2.928	0.001
Phytochemicals	--	--	--	--
Geographical position	MEM13 MEM14	0.182	2.778	0.002

BIBLIOGRAPHIE (ARTICLE 2)

AFSQ (2018) La clé forestière: Le frêne blanc. https://afsq.org/information-foret/nos-arbres/frene-blanc/.

Agler MT, Ruhe J, Kroll S, Morhenn C, Kim S-T, Weigel D, Kemen EM (2016) Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS biology* 14(1):e1002352.

Aukema JE, Leung B, Kovacs K, Chivers C, Britton KO, Englin J, Frankel SJ, Haight RG, Holmes TP, Liebhold AM (2011) Economic impacts of non-native forest insects in the continental United States. *PLoS One* 6(9):e24587.

Beauchemin NJ, Furnholm T, Lavenus J, Svistoonoff S, Doumas P, Bogusz D, Laplaze L, Tisa LS (2012) Casuarina root exudates alter the physiology, surface properties, and plant infectivity of Frankia sp. strain CcI3. *Appl. Environ. Microbiol.* 78(2):575-580.

Bi JL, Toscano NC, Madore MA (2003) Effect of Urea Fertilizer Application on Soluble Protein and Free Amino Acid Content of Cotton Petioles in Relation to Silverleaf Whitefly (Bemisia argentifolii) Populations. *Journal of Chemical Ecology* 29(3):747-761.

Biedermann PH & Vega FE (2020) Ecology and evolution of insect–fungus mutualisms. *Annual Review of Entomology* 65:431-455.

Blanchet FG, Legendre P, Borcard D (2008) Forward selection of explanatory variables. *Ecology* 89(9):2623-2632.

Borcard D, Gillet F, Legendre P (2018) *Numerical ecology with R.* Springer,

Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology* 73(3):1045-1055.

Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal* 11(12):2639.

Cappaert D, McCullough DG, Poland TM, Siegert NW (2005a) Emerald ash borer in North America: a research and regulatory challenge. *American Entomologist. 51 (3): 152-165.* 51(3).

Chen Y & Poland TM (2009) Interactive influence of leaf age, light intensity, and girdling on green ash foliar chemistry and emerald ash borer development. *J Chem Ecol* 35(7):806-815.

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42(Database issue):D633-642.

Colman DR, Toolson EC, Takacs-Vesbach C (2012) Do diet and taxonomy influence insect gut bacterial communities? *Molecular ecology* 21(20):5124-5137.

Comeau AM, Li WKW, Tremblay J-É, Carmack EC, Lovejoy C (2011) Arctic Ocean Microbial Community Structure before and after the 2007 Record Sea Ice Minimum. *PLOS ONE* 6(11):e27492.

De Cáceres M & Jansen F (2020) Package 'indicspecies': relationship between species and groups of sites. R Package Version 1.7. 9.

De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* 119(10):1674-1684.

Douglas AE (2015) Multiorganismal insects: diversity and function of resident microorganisms. *Annual review of entomology* 60:17-34.

Dray S, Bauman D, Blanchet G, Borcard D, Clappe S, Guenard G (2020) & Wagner, HH (2020). adespatial: Multivariate multiscale spatial analysis (R package version 0.3-8).).

Edgar RC (2010b) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19):2460-2461.

Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv* :081257.

Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biological conservation* 61(1):1-10.

Feldhaar H (2011) Bacterial symbionts as mediators of ecologically important traits of insect hosts. *Ecological Entomology* 36(5):533-543.

Ferguson LV, Dhakal P, Lebenzon JE, Heinrichs DE, Bucking C, Sinclair BJ (2018) Seasonal shifts in the insect gut microbiome are concurrent with changes in cold tolerance and immunity. *Functional ecology* 32(10):2357-2368.

Fischer R, Ostafe R, Twyman RM (2013) *Yellow Biotechnology II:Insect Biotechnology in Plant Protection and Industry,* Springer. p 51-64.

Franzini PZ, Ramond J-B, Scholtz CH, Sole CL, Ronca S, Cowan DA (2016) The gut microbiomes of two Pachysoma MacLeay desert dung beetle species (Coleoptera: Scarabaeidae: Scarabaeinae) feeding on different diets. *PloS one* 11(8):e0161118.

Gandhi KJ & Herms DA (2010a) North American arthropods at risk due to widespread Fraxinus mortality caused by the alien emerald ash borer. *Biological Invasions* 12(6):1839-1846.

García AMn, Moumen A, Ruiz DY, Alcaide EM (2003) Chemical composition and nutrients availability for goats and sheep of two-stage olive cake and olive leaves. *Animal Feed Science and Technology* 107(1-4):61-74.

Garg B (2012) *Plant Analysis: Comprehensive Methods and Protocols.* Scientific Publishers,

Glassman SI & Martiny JBH (2018) Broadscale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. *mSphere* 3(4):e00148-00118.

Gupta A & Nair S (2020) Dynamics of Insect–Microbiome Interaction Influence Host and Microbial Symbiont. *Frontiers in microbiology* 11:1357.

Hagerman AE (1988) Extraction of tannin from fresh and preserved leaves. *Journal of chemical Ecology* 14(2):453-461.

Herms DA & McCullough DG (2014) Emerald ash borer invasion of North America: history, biology, ecology, impacts, and management. *Annu Rev Entomol* 59:13-30.

Howe GA & Schaller A (2008) Direct defenses in plants and their induction by wounding and insect herbivores. *Induced plant resistance to herbivory,* Springer. p 7-29.

Humphrey PT & Whiteman NK (2020) Insect herbivory reshapes a native leaf microbiome. *Nature Ecology & Evolution* 4(2):221-229.

Kadivar H & Stapleton AE (2003) Ultraviolet radiation alters maize phyllosphere bacterial diversity. *Microbial Ecology* 45(4):353-361.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26(11):1463-1464.

Kembel SW, Eisen JA, Pollard KS, Green JL (2011) The phylogenetic diversity of metagenomes. *PLoS One* 6(8):e23214.

Kembel SW, O'Connor TK, Arnold HK, Hubbell SP, Wright SJ, Green JL (2014) Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proceedings of the National Academy of Sciences* 111(38):13715-13720.

Kim JM, Choi M-Y, Kim J-W, Lee SA, Ahn J-H, Song J, Kim S-H, Weon H-Y (2017) Effects of diet type, developmental stage, and gut compartment in the gut bacterial communities of two Cerambycidae species (Coleoptera). *Journal of Microbiology* 55(1):21-30.

Kovacs KF, Haight RG, McCullough DG, Mercader RJ, Siegert NW, Liebhold AM (2010) Cost of potential emerald ash borer damage in U.S. communities, 2009–2019. *Ecological Economics* 69(3):569-578.

Legendre P & De Cáceres M (2013) Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology letters* 16(8):951-963.

Lennon JT, Muscarella ME, Placella SA, Lehmkuhl BK (2018) How, When, and Where Relic DNA Affects Microbial Diversity. *mBio* 9(3):e00637-00618.

Marquis RJ, Newell EA, Villegas AC (1997) Non-structural carbohydrate accumulation and use in an understorey rain-forest shrub and relevance for the impact of leaf herbivory. *Functional Ecology* 11(5):636-643.

Mason CJ, Hanshew AS, Raffa KF (2015) Contributions by Host Trees and Insect Activity to Bacterial Communities in Dendroctonus valens (Coleoptera: Curculionidae) Galleries, and Their High Overlap With Other Microbial Assemblages of Bark Beetles. *Environmental Entomology* 45(2):348-356.

McCullough DG (2020) Challenges, tactics and integrated management of emerald ash borer in North America. *Forestry: An International Journal of Forest Research* 93(2):197-211.

Mittapalli O, Bai X, Mamidala P, Rajarapu SP, Bonello P, Herms DA (2010) Tissue-specific transcriptomics of the exotic invasive insect pest emerald ash borer (Agrilus planipennis). *PLoS one* 5(10):e13708.

Mogouong J, Constant P, Lavallée R, Guertin C (2020) Gut microbiome of the emerald ash borer, Agrilus planipennis Fairmaire, and its relationship with insect population density. *FEMS Microbiology Ecology* 10.1093/femsec/fiaa141.

Moran NA & Yun Y (2015) Experimental replacement of an obligate insect symbiont. *Proceedings of the National Academy of Sciences* 112(7):2093-2096.

Oksanen J, Blanchet F, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin P, O'Hara R, Simpson G, Solymos P (2019) Package 'vegan'. R package version 2.5-6.

Oliverio AM, Gan H, Wickings K, Fierer N (2018) A DNA metabarcoding approach to characterize soil arthropod communities. *Soil Biology and Biochemistry* 125:37-43.

Peres-Neto PR, Legendre P, Dray S, Borcard D (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* 87(10):2614-2625.

Poland TM & McCullough DG (2006) Emerald ash borer: invasion of the urban forest and the threat to North America's ash resource. *Journal of Forestry* 104(3):118-124.

Popa V, Deziel E, Lavallee R, Bauce E, Guertin C (2012) The complex symbiotic relationships of bark beetles with microorganisms: a potential practical approach for biological control in forestry. *Pest Manag Sci* 68(7):963-975.

Qadri M, Short S, Gast K, Hernandez J, Wong AC-N (2020) Microbiome Innovation in Agriculture: Development of Microbial Based Tools for Insect Pest Management. *Front. Sustain. Food Syst. 4: 547751. doi: 10.3389/fsufs*.

Schowalter TD (2016) *Insect ecology: an ecosystem approach.* Academic press,

Shikano I, Rosa C, Tan C-W, Felton GW (2017) Tritrophic interactions: microbe-mediated plant effects on insect herbivores. *Annual review of Phytopathology* 55:313-331.

Slesak RA, Lenhart CF, Brooks KN, D'Amato AW, Palik BJ (2014) Water table response to harvesting and simulated emerald ash borer mortality in black ash wetlands in Minnesota, USA. *Canadian Journal of Forest Research* 44(8):961-968.

Smith CC, Snowberg LK, Caporaso JG, Knight R, Bolnick DI (2015) Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *The ISME journal* 9(11):2515-2526.

Stam JM, Kroes A, Li Y, Gols R, van Loon JJ, Poelman EH, Dicke M (2014) Plant interactions with multiple insect herbivores: from community to genes. *Annual review of plant biology* 65:689-713.

Thapa S & Prasanna R (2018) Prospecting the characteristics and significance of the phyllosphere microbiome. *Annals of microbiology* 68(5):229-245.

Toju H, Tanabe AS, Yamamoto S, Sato H (2012) High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PloS one* 7(7):e40863.

Torti SD, Dearing MD, Kursar TA (1995) Extraction of phenolic compounds from fresh leaves: a comparison of methods. *Journal of Chemical Ecology* 21(2):117-125.

Vacheron J, Péchy-Tarr M, Brochet S, Heiman CM, Stojiljkovic M, Maurhofer M, Keel C (2019) T6SS contributes to gut microbiome invasion and killing of an herbivorous pest insect by plant-beneficial Pseudomonas protegens. *The ISME Journal* 13(5):1318-1329.

Van Soest Pv, Robertson J, Lewis B (1991) Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *Journal of dairy science* 74(10):3583-3597.

Vasanthakumar A, Handelsman J, Schloss PD, Bauer LS, Raffa KF (2008) Gut microbiota of an invasive subcortical beetle, Agrilus planipennis Fairmaire, across various life stages. *Environmental Entomology* 37(5):1344-1353.

Vorholt JA (2012) Microbial life in the phyllosphere. *Nature Reviews Microbiology* 10(12):828.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology* 73(16):5261-5267.

Watanabe H & Tokuda G (2010) Cellulolytic systems in insects. *Annual review of entomology* 55.

Wellburn R (1994) The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *Journal of plant physiology* 144(3):307-313.

Wielkopolan B & Obrepalska-Steplowska A (2016) Three-way interaction among plants, bacteria, and coleopteran insects. *Planta* 244(2):313-332.

Yutthammo C, Thongthammachat N, Pinphanichakarn P, Luepromchai E (2010) Diversity and activity of PAH-degrading bacteria in the phyllosphere of ornamental plants. *Microbial ecology* 59(2):357-368.

Zhang Z, Jiao S, Li X, Li M (2018) Bacterial and fungal gut communities of Agrilus mali at different developmental stages and fed different diets. *Scientific Reports* 8(1):15634.

3 ANALYSE DE LA COMPLEXITE DU MICROBIOME DU TRACT INTESTINAL DE L'AGRILE DU FRENE ADULTE : STRUCTURE PHYLOGÉNÉTIQUE ET RESEAUX DE COOCCURRENCES

Title in english : The complexity of the Emerald Ash Borer gut microbiome: phylogenetic structure and cooccurrences networks

Lien entre l'article ou les articles précédents et le suivant :

La plasticité du microbiome intestinal vis-à-vis de la densité de la population au sein des arbres hôtes a mis en évidence une relation significative entre le microbiome de l'agrile du frêne et certains paramètres intrinsèques à l'arbre hôte, notamment ceux associés aux mécanismes de défense suivant l'attaque par un insecte herbivore. Le deuxième article qui cherchait à élucider les effets potentiels de quatre propriétés intrinsèques aux arbres hôtes (phytochimie des feuilles, communautés bactériennes et fongiques et leurs positions géographiques) sur le microbiome de l'insecte sur la base du partitionnement de la variation observée, a suggéré un rôle significatif des communautés microbiennes associées à la phyllosphère des arbres hôtes. Ces résultats ont révélé d'une part l'effet de pression sélective des deux biotopes que sont le tract intestinal de l'insecte et la phyllosphère, d'autre part des interactions potentielles entre les ASV des communautés microbiennes notamment avec certains ASV indicateurs écologiques de chaque biotope. Cet article présente une analyse exploratoire comparative basée sur l'analyse de la parenté phylogénétique des communautés microbiennes et les cooccurrences des taxa dans chaque hôte. Les résultats de cet article seront soumis pour une publication, mais ils apportent des éléments importants qui devraient permettre d'envisager de nouvelles avenues de recherche pour une meilleure compréhension théorique des processus écologiques d'assemblage inhérents aux communautés microbiennes au sein du tractus intestinal d'un insecte ravageur comme l'agrile du frêne.

3.1 Abstract

Microorganisms associated with the gut have been extensively documented for the benefits they confer to their host. These microorganisms are reported to be relatively diverse, with a variety of functional properties, including functions related to biology, ecology, or host evolution. Several factors are now identified as having a modulating role in the structure of the microbiome in its habitat. Among these modulating factors, microbe-microbe interactions are reported as critical to the health of microbiota in both animals and plants, and above all, they would contribute to equilibrium in host interactions with its microbiome. Our previous results showing variations in the adult EAB gut microbiome under the pressure of the microbiome of the phyllosphere had highlighted the important place of complex microbe-microbe relationships in the insect's gut microbiome structure. In this chapter, a concept based on the analysis of the phylogenetic structure and species co-occurrences revealed, respectively, the determining effect of the host on the assemblage of the microbial species, as well as the complexity of the structure of the microbiome structure in the insect's gut compared to the phyllosphere.

3.2 Résumé

Les microorganismes associés au tract intestinal ont été largement documentés pour leurs bienfaits pour leur hôte. Ces microorganismes sont rapportés comme étant relativement diversifiés, avec des propriétés fonctionnelles variées incluant les fonctions en lien avec la biologie, l'écologie ou encore l'évolution de l'hôte. Plusieurs facteurs sont aujourd'hui identifiés comme ayant un rôle modulant la structure du microbiome dans son habitat. Parmi ces facteurs modulants, les interactions microbe-microbe sont rapportées comme étant critique pour les fonctions du microbiote aussi bien chez les animaux que chez les plantes, et surtout elles contribueraient à l'équilibre dans les interactions de l'hôte avec son microbiome. Nos précédents résultats montrant des variations du microbiome intestinal de l'insecte adulte sous la pression du microbiome de la phyllosphère avait mis en évidence la place importante des relations complexes microbe-microbe dans le tract intestinal de l'insecte. Dans ce travail, un concept basé sur l'analyse de la structure phylogénétique et des cooccurrences des espèces a permis de mettre en évidence l'effet déterminant du biotope (hôte) sur l'assemblage des espèces, ainsi que quelques éléments de la complexité de la structure du microbiome dans le tract intestinal de l'insecte comparativement à la phyllosphère.

3.3 Introduction

Les bienfaits des microorganismes associés au tract intestinal exercent plusieurs bénéfices pour leur hôte. De plus ces microorganismes sont rapportés comme étant relativement diversifiés leur conférant des propriétés fonctionnelles variées incluant les fonctions en lien avec la biologie, l'écologie ou encore l'évolution de l'hôte. Plusieurs facteurs sont aujourd'hui identifiés comme jouant un rôle modulant sur la structure du microbiome, notamment le régime alimentaire, le génotype de l'hôte et des facteurs environnements reliés affectant l'état physiologique de l'hôte. Parmi ces déterminants, les interactions microbe-microbe sont rapportées comme étant critique pour la santé du microbiote aussi bien chez les animaux que chez les plantes, et surtout elles contribueraient à l'équilibre dans les interactions de l'hôte avec son microbiome (Simon *et al.*, 2019). Par exemple, les travaux de Hunter *et al.* (2010) ont démontré une interaction significative entre deux espèces du genre *Erwinia* sp. et du genre *Enterobacter* sp., prédominantes sur la phyllosphère de la laitue (*Lactuca* species) et dont la présence a été rapportée comme étant sensible au niveau de contamination par le genre *Salmonella* ou encore l'espèce *E. coli* O157 :H7. De plus les travaux récents de Itoh *et al.* (2019) suggèrent que la compétition microbe-microbe a un impact sur l'évolution et la stabilité des associations symbiotiques maintenues par transmission de symbiontes environnementaux dans le tract intestinal des insectes. Nos précédents résultats montrant des variations du microbiome intestinal des adultes d'*Agrilus planipennis* sous l'influence du microbiome de la phyllosphère avaient aussi mis en évidence la place importante des relations complexes microbe-microbe dans la stabilité de la structure du microbiome du tract intestinal. Ces résultats amènent à se questionner sur les processus écologiques d'assemblage des communautés qui cooccurrent dans le tract intestinal de l'insecte. Bien que les microorganismes associés à la phyllosphère des arbres hôtes aient été révélés ayant un effet déterminant sur la structure taxonomique du microbiome du tract intestinal des agriles adultes (Mogouong *et al.*, 2020; Mogouong *et al.*, 2021), d'autres variables pourraient jouer un rôle déterminant sur cette structure, notamment l'information phylogénétique des communautés. La diversité phylogénétique (PD) proposée par (Faith, 1992) qui mesure les distances évolutionnaires totales à travers les espèces dans une communauté s'était avérée plus faible dans le tract intestinal de l'insecte pour la communauté fongique. D'autres approches proposées permettent de mieux caractériser cette variable en évaluant la structure phylogénétique d'une communauté d'espèces à travers l'analyse de la « surdispersion » phylogénétique (espèces dispersées sur la base de leur phylogénie) ou encore du « regroupement » phylogénétique (espèces rapprochées sur la base de leur phylogénie) (Kembel *et al.*, 2010; Kembel & Hubbell,

2006; Webb *et al.*, 2002). Les travaux de Webb *et al.* (2002) ont établi un lien de causalité entre la structure phylogénétique d'une communauté d'espèces et l'effet de son environnement sur le processus d'assemblage de ces espèces dans leur habitat. Cette relation de causalité a déjà été rapportée chez d'autres espèces, notamment dans un modèle aviaire (van Veelen *et al.*, 2017), un modèle coléoptère (Vamosi & Vamosi, 2007), ou encore chez des acariens (Gong *et al.*, 2018). Dans tous ces cas, il est rapporté que le regroupement phylogénétique est déterminé par un effet significatif de l'espace habitat sur les taxa qui y cooccurrent. La surdispersion phylogénétique des taxa dans des communautés naturelles associées à différents habitats (sols, milieu aquatique, et sédiments) pourrait être due à des capacités de dispersion et/ou de colonisation, ou encore à un événement de rayonnement adaptatif (Horner-Devine & Bohannan, 2006). On peut supposer que l'évaluation de la structure phylogénétique du microbiome du tract intestinal des adultes de l'agrile du frêne pourrait permettre d'évaluer l'importance de l'effet de l'habitat sur l'assemblage des espèces qui y cooccurrent. Bien que l'intérêt pour la prise en compte de la composante phylogénétique dans les études de microbiome soit en progression depuis les dernières décennies, les connaissances restent limitées à ce jour pour ce qui est du microbiome des insectes. Par exemple, les travaux de Gong *et al.* (2018) ont révélé dans leur étude que les communautés bactériennes et fongiques du tract intestinal de plusieurs types d'insectes étaient caractérisées par des espèces étroitement apparentées (structure phylogénétique ayant une tendance au regroupement) suggérant ainsi un effet déterminant du filtrage du biotope que représente le tract intestinal sur le microbiome de ces insectes. Une analyse exploratoire dans le cadre de nos précédentes analyses a permis de mettre en évidence une diversité phylogénétique (PD) comparable pour la communauté bactérienne dans les deux biotopes (représentés ici par les deux biotopes que sont le tract intestinal de l'insecte et la phyllosphère), contrairement à la communauté fongique qui indiquait une PD plus faible chez l'insecte. La littérature rapporte que la diversité PD n'intègre pas toute l'information sur l'abondance relative des espèces au sein d'une communauté (Cadotte *et al.*, 2010). Le tract intestinal est présenté dans la littérature comme étant un environnement présentant des conditions assez rudes pour la croissance de microorganismes. Dillon and Dillon (2004) ont par exemple suggéré que de nombreuses catégories d'insectes avaient des tracts intestinaux compartimentés supposant l'existence dans cet environnement de plusieurs niches spécialisées. De plus, le tract intestinal des insectes est moins exposé à certains facteurs environnements tels que les rayons ultraviolets ou encore la pollution comme peut l'être la phyllosphère. Les travaux de Yun *et al.* (2014), qui comparaient les populations bactériennes à travers le tract intestinal chez plusieurs espèces d'insectes, avaient suggéré que la dynamique de la variation du microbiote du tract intestinal pourrait être déterminée par la morphologie du

tract et ses conditions physicochimiques. Dans l'article 2 de cette thèse, l'effet sélectif/filtrant du biotope sur le microbiome a été confirmé avec l'identification de taxas indicateurs et la mise en évidence des relations microbes-microbes complexes dans chaque biotope. Bien que de plus en plus d'études s'y intéressent aux relations microbes-microbes, il n'existe actuellement pas de données disponibles se rapportant à notre modèle d'étude. Ainsi, la représentation des interactions microbe-microbe basée sur l'analyse des réseaux écologiques pourrait aider à mieux comprendre le rôle des interactions dans la coexistence des espèces dans un habitat et la stabilité des structures taxonomiques formées par l'assemblage de ces espèces (van der Heijden & Hartmann, 2016). Une analyse des réseaux dans chaque biotope permet de révéler des éléments tels que la cooccurrence ou la co-exclusion de certains taxa microbiens. Ainsi, l'analyse des réseaux de co-occurrences pourrait permettre d'évaluer la complexité des relations microbe-microbe dans chaque biotope, tout en assumant que ces cooccurrences peuvent être modulées par des pressions exercées par des contraintes liées au biotope ou d'autres paramètres du microenvironnement. Les modèles de cooccurrences ont été explorés entre les communautés microbiennes de chaque biotope en utilisant l'approche probabiliste proposée par Griffith *et al.* (2016) et adapté par Veech (2013). Cette approche est basée sur l'utilisation de fréquences observées de cooccurrences entre chaque paire de taxa et la distribution de chaque taxon. Ces fréquences observées sont ensuite utilisées pour calculer les fréquences attendues respectives et d'en déduire les probabilités que deux espèces apparaissent à une fréquence plus élevée (ou plus faible) que la fréquence d'apparition observée. Ainsi l'algorithme permet de déterminer une fréquence d'occurrence observée comme étant une association positive, négative ou aléatoire. Cette démarche favorise la mise en évidence de certains taxa potentiellement importants dans l'assemblage des espèces cooccurrentes. Les travaux proposés dans cet article visent à explorer la structure phylogénétique du microbiome et les cooccurrences de taxa peuvent apporter des éléments de compréhension complémentaires sur la dynamique du microbiome dans son environnement. Pour cela, les données de séquençage d'amplicons issues de l'ADN du tract intestinal de l'agrile adulte et de la phyllosphère ont permis d'étudier la structure phylogénétique des communautés bactérienne et fongique associées à chaque biotope en utilisant l'approche basée sur la conservation des caractéristiques écologiques phylogénétiques à partir de l'arbre phylogénétique proposée par Webb *et al.* (2002), et adapté dans des travaux scientifiques publiés (Kembel *et al.*, 2010; Kembel & Hubbell, 2006). Enfin, une analyse des cooccurrences d'espèces a été effectuée pour déterminer les cooccurrences de taxa significativement positives ou négatives utilisant une approche probabiliste basée sur un modèle nul de permutation (Griffith *et al.*, 2016).

3.4 Matériel et méthodes

3.4.1 Localisation des sites d'échantillonnage, collecte des insectes et des feuilles de frêne

Vingt-neuf arbres (frênes) ont été identifiés et la collecte des insectes et des feuilles a été effectuée en suivant les étapes décrites dans l'article 2 (Mogouong *et al.*, accepté pour publication dans le journal *Scientific Reports*).

3.4.2 Extraction de l'ADN, séquençage haut débit d'amplicons et traitement bio-informatique des données de séquençage

L'extraction de l'ADN et les conditions de séquençages ont été effectuées en suivant les étapes décrites dans l'article 2 de cette thèse. Les données finales correspondaient à une table d'ASV bactérien comprenant 518 ASV représentés par 410 125 séquences et une table d'ASV fongiques comprenant 281 ASV représentés par 2 281 382 séquences.

3.4.3 Analyses statistiques

Structure phylogénétique des communautés microbiennes : Dans notre approche basée sur les travaux de Kembel and Hubbell (2006), deux métriques ont été utilisées : la moyenne de la distance phylogénétique observée ou « Mean pairwise distance » (MPD) correspondant à la parenté des espèces à la base des branches dans de l'arbre phylogénétique (patron global de la structure phylogénétique) et la distance moyenne du taxon le plus proche observée ou « Mean nearest taxon distance » (MNTD) correspondant à la parenté aux extrémités des branches dans l'arbre phylogénétique (patron au niveau des clades) (Webb, 2000). Finalement, ces deux mesures sont évaluées à partir de deux indices spécifiques, l'indice de parenté nette (NRI) pour la MPD et l'indice de taxon le plus proche (NTI) pour la MNTD. Le NRI permet de mesurer le regroupement ou la surdispersion phylogénétique des espèces à l'échelle de l'arbre, tandis que le NTI permet de mesurer le regroupement ou la surdispersion phylogénétique terminale (Webb, 2000). La librairie *Picante* Version:1.8.1 (Kembel *et al.*, 2010) a été utilisée pour calculer la MNTD observée ($MNTD_{obs}$) et la MPD observée (MPD_{obs}) parmi toutes les paires d'espèces de chaque communauté. Les valeurs observées ont ensuite permis de générer des distances phylogénétiques moyennes standardisées ou « standardized effect size mean phylogenetic distances » ou SES_{MPD} (ou SES_{MNTD}) en utilisant la fonction ses.mpd() (ou ses.mntd()) qui se base sur un modèle nul utilisant une probabilité égale à partir de la matrice de distance phylogénétique pour chaque espèce ("phylogeny.pool") avec 999 permutations (Gong *et al.*, 2018) et qui

correspondent à des mesures de la diversité des traits phylogénétiques à travers les communautés (plus la valeur est élevée, plus les espèces cooccurrentes sont distantes phylogénétiquement). Enfin, le NRI est obtenu en multipliant par -1 le SES_{MPD} et les NTI en multipliant par -1 le SES_{MNTD}. Les valeurs négatives du NRI (NTI) indiquent la dispersion phylogénétique (les taxa qui cooccurrent ont une très faible relation de parenté comparée à celle attendue par le modèle nul), tandis que les valeurs positives du NRI (NTI) indiquent un regroupement phylogénétique (les taxa qui cooccurrent ayant une parenté plus proche que celle attendue dans le modèle nul). Des valeurs du NRI positives indiquent que les espèces qui cooccurrent sont plus étroitement apparentées phylogénétiquement que le fruit du hasard, généralement à cause du regroupement observé sur l'arbre phylogénétique. Tandis que des valeurs du NTI positives indiquent que les espèces qui cooccurrent ont tendance à se retrouver avec des espèces étroitement apparentées (congénères), généralement à cause du regroupement phylogénétique terminal. À l'inverse, les valeurs négatives du NRI révèlent que les espèces qui cooccurrent sont moins apparentées, ce qui devrait se traduire par une surdispersion basée sur l'arbre phylogénétique, alors que les valeurs négatives de NTI sont une indication que les espèces cooccurrentes n'ont pas tendance à se regrouper avec d'autres espèces entrainant une surdispersion au niveau des clades. Un aspect plus détaillé de la structure phylogénétique a été évalué en observant le ratio RNI/NTI qui indique si les échantillons se trouvent dans un nombre restreint de groupes phylogénétiques (ratio NRI/NTI élevé) ou dans plusieurs groupes phylogénétiques (ratio NRI/NTI faible) (Carson & Schnitzer, 2011). Ces analyses ont permis d'évaluer le degré de regroupement ou de surdispersion phylogénétique dans chaque échantillon en relation avec le modèle nul de permutation. Par la suite des analyses en coordonnées principales ont été effectuées en utilisant successivement les distances Bray-Curtis, MTD et MNPD afin d'évaluer la finesse apportée par les données relatives à la structure phylogénétique sur l'analyse de la bétadiversité entre ces deux biotopes.

Analyse des réseaux de cooccurrences : Le modèle d'analyse des réseaux de cooccurrences est basé sur le calcul de la probabilité P_j des espèces 1 et 2 d'apparaître exactement sur un nombre de sites j (les valeurs de j varient de 0 à N). N correspond au nombre de sites (feuilles/insectes). Pour un nombre de sites donné N et un ensemble de sites j, il existe un nombre limité de possibilités que les espèces 1 et 2 peuvent être distribuées à travers les sites de manière à maintenir les nombres de sites observés N_1 et N_2. Ainsi, le modèle probabiliste utilisé ici correspond à déterminer le nombre de possibilités r éléments parmi l'ensemble composé de n éléments symbolisés par un calcul de combinaisons C(n,r) (Veech, 2013).

Ainsi, les cooccurrences ont été basées sur deux éléments : la probabilité que deux espèces apparaissent à une fréquence plus faible que la fréquence moyenne d'apparition P_{lt} et la probabilité que deux espèces apparaissent à une fréquence plus élevée que la fréquence d'apparition observée P_{gt}. Si ces deux probabilités sont inférieures à 0.05, cela impliquerait des cooccurrences négatives (potentiellement une interaction négative) ou des cooccurrences positives lorsque les deux probabilités sont supérieures à 0.05. Dans cette analyse, le modèle met aussi en évidence de nombreuses associations aléatoires d'espèces identifiées par une faible puissance statistique (Griffith *et al.*, 2016; Veech, 2013).

$$P_{lt} = \sum P_j \, pour \, j = 0 \text{ à } Q_{Obs} - 1$$

$$P_{gt} = \sum P_j \, pour \, j = Q_{Obs} + 1 \text{ à } N$$

Q_{Obs} correspond à la fréquence des cooccurrences observées et P_j la probabilité que deux espèces apparaissent exactement à un site spécifique j (les valeurs de j varient de 0 à N). Les matrices utilisées contenaient les valeurs binaires de présence ou absence, pour chaque biotope (feuille et insecte), d'ASV des communautés bactérienne et fongique ayant une occurrence dans au moins 30% des échantillons. Les données ont ensuite été transformées pour ne considérer que des valeurs de présence-absence. La librairie *cooccur* Version: 1.3 (Griffith *et al.*, 2016) a été utilisée pour calculer les probabilités et la représentation des réseaux a été effectuée avec la fonction ggnet2() de la librairie *ggnet* (Briatte, 2021). Finalement, les réseaux obtenus ont été évalués en se basant sur deux unités de mesure, la modularité et la présence d'ASV considéré comme des taxa « clés » sur la base de leur effet concentrateur de connexions (Agler *et al.*, 2016). La modularité permet de mesurer le partitionnement le plus approprié d'un réseau de connexions en se basant sur le fait que dans un réseau, il y a de nombreuses connexions entre les espèces de certaines communautés et seulement quelques connexions entre ces communautés (Clauset *et al.*, 2004; Newman & Girvan, 2004). Elle caractérise l'existence de sous-réseaux de nœuds dans lesquels la densité des interactions est plus élevée à l'intérieur des sous-réseaux que celle entre les sous-réseaux (Layeghifard *et al.*, 2017; Olesen *et al.*, 2007).

Le réseau écologique est défini par un ensemble de nœuds (les acteurs) et un ensemble de connexions (liens entre les acteurs). Dans chaque réseau, il existe une distance entre les acteurs, et cette valeur est nulle entre un acteur et lui-même et peut être variable d'un acteur à un autre. Deux métriques topologiques de connexion ont été calculées afin de caractériser et de comparer les réseaux écologiques identifiés. Dans un premier temps, le degré de nœud correspondant à

une mesure quantitative qui permet de savoir à quel point un acteur (ASV) est connecté et correspond ici au nombre de relations de cooccurrences dans laquelle est impliqué un ASV dans son réseau écologique (Montoya et al., 2006), il est corrélé à la densité du réseau. Ensuite, la centralité de proximité d'un ASV ou « closeness centrality » correspondant à la proximité d'un ASV avec tous les autres ASV du réseau, les valeurs élevées indiquent les ASV concernés peuvent facilement être en relation avec tous les autres ASV du réseau (Freeman *et al.*, 1979).

$$CC_i = \sum_{j-1;\, i \neq j}^{n} \frac{\mathrm{d_{ij}}}{\mathrm{n} - 1}$$

CC_i = centralité de proximité de l'espèce i

n = est le nombre d'espèces

$\mathrm{d_{ij}}$ = distance la plus courte entre les espèces i et j mesurée en nombre de liens

Les deux métriques ont été calculées en utilisant des fonctions se trouvant dans la librairie iGraph, respectivement *degree() et closeness()* proposées par Csardi and Nepusz (2006). L'utilisation de ces métriques a permis de mettre en évidence certains ASV fortement connectés et qui pourraient jouer un rôle clé dans la stabilité des réseaux écologiques respectifs.

Certains travaux ont mis en évidence l'importance de certains nœuds dans un réseau écologique en se basant sur des métriques topologiques de connectivité (Banerjee *et al.*, 2016; González *et al.*, 2010). Ainsi, la notion d'espèces (ASV) clés ou concentrateurs en raison de leur position dans un réseau de connexions a été suggérée pour les espèces ayant des valeurs de degré de nœud et de centralité de proximité élevées. Ces taxa présentant un nombre particulièrement élevé de connexions pourraient jouer un rôle disproportionné dans la formation du réseau et dans la communauté.

3.5 Résultats

3.5.1 **Structure phylogénétique des communautés microbiennes**

En se basant sur l'indice de diversité PD, la communauté bactérienne est apparue avec une diversité phylogénétique proche dans les deux biotopes contrairement à la communauté fongique dans le tract intestinal des insectes qui a révélé une différence significative entre les deux biotopes (**Figure 3.1**). Cependant, les paramètres relatifs à la structure phylogénétique ont permis de mettre en évidence beaucoup plus de contraste entre les deux biotopes. Les communautés bactériennes tout comme les communautés fongiques sont apparues avec des ASV plus

étroitement apparentés que prévu par hasard (SES(MPD) et SES(MNTD) inférieurs à zéro). Cependant, la communauté fongique du tract intestinal a révélé des ASV beaucoup moins étroitement apparentés que celle de la phyllosphère. Les résultats obtenus suggèrent une différenciation écologique (phylogénétique) plus élevée au niveau du tract intestinal de l'insecte, notamment pour la communauté fongique.

Le ratio RNI/NTI de la communauté bactérienne était significativement plus élevé dans le tractus intestinal de l'insecte qu'au niveau de la phyllosphère avec respectivement des valeurs moyennes de 0,31 et 0,16 suggérant une tendance à former moins de regroupement phylogénétique au niveau du tract intestinal de l'insecte et donc une différenciation écologique précoce (**Tableau Suppl. 4.1**). À l'inverse, le ratio NRI/NTI de la communauté fongique était beaucoup plus faible au niveau du tract intestinal de l'insecte qu'au niveau de la phyllosphère avec respectivement une valeur moyenne -2,45 et +0,27 suggérant une tendance pour cette communauté à former plus de regroupement au niveau du tract intestinal de l'insecte, et donc une différenciation écologique plus tardive que dans le cas de la communauté bactérienne.

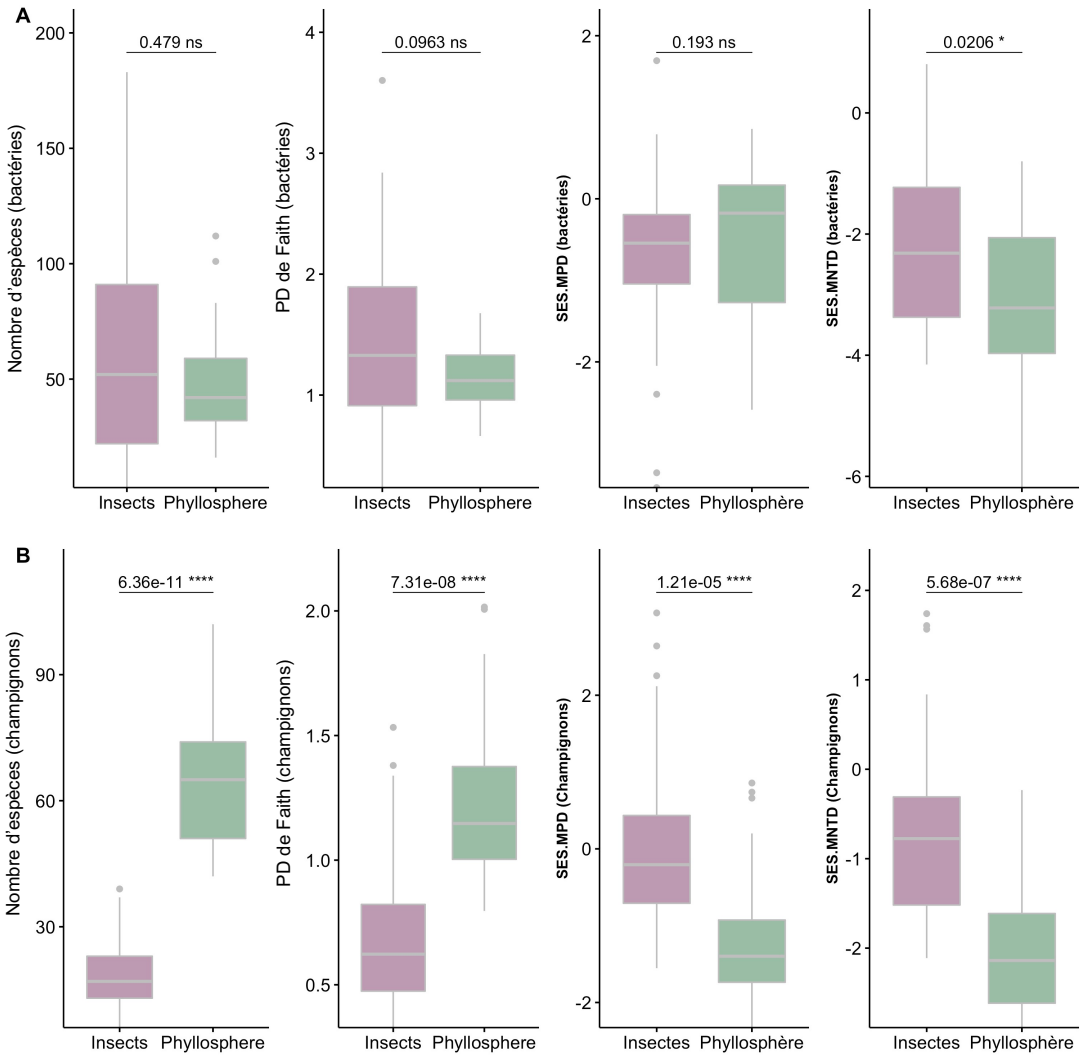


Figure 3.1 **Comparaison des mesures de diversité et structure phylogénétique dans les deux biotopes (n=29 par biotope) pour les communautés bactérienne (A) et fongique (B)**

3.5.2 Analyse de la bétadiversité

La bétadiversité basée sur les distances Bray-Curtis nous indique un regroupement significatif pour chacune des deux communautés, confirmant ainsi nos précédents résultats (*p*-value = 0.001), **Figure 3.2**. En revanche, la communauté bactérienne est apparue avec un pourcentage de variation (expliquées par les deux premiers axes) plus élevé en se basant sur les traits phylogénétiques (distance MNTD) comparativement à la distance Bray-Curtis (non phylogénétique) respectivement 16% et 72% (Figure 3.2). Enfin, la communauté fongique a montré deux groupes distincts en fonction de leurs biotopes basés sur la distance Bray-Curtis (*p*-value = 0.034) contrairement à la distance phylogénétique (*p*-value = 0.0507), **Figure 3.3.**

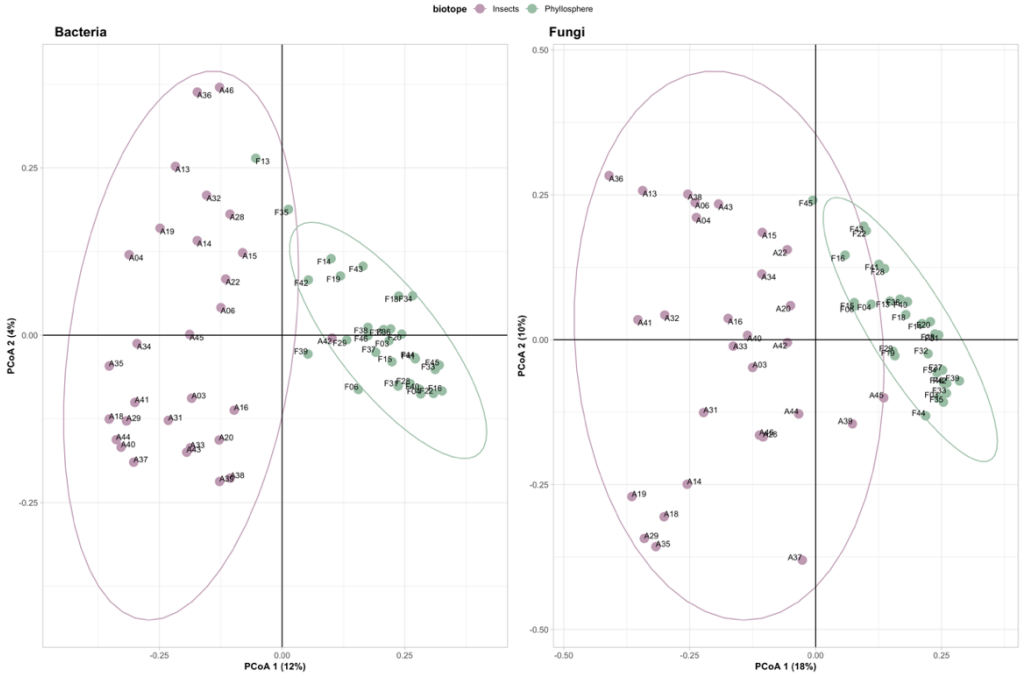


Figure 3.2 PCoA basée sur les distances Bray-Curtis des données d'abondance de la communauté bactérienne, p-value < 0.001 (à gauche) et la communauté fongique, p-value < 0.001 (droite). Pour chaque biotope, le nombre d'échantillons considérés est de 29.

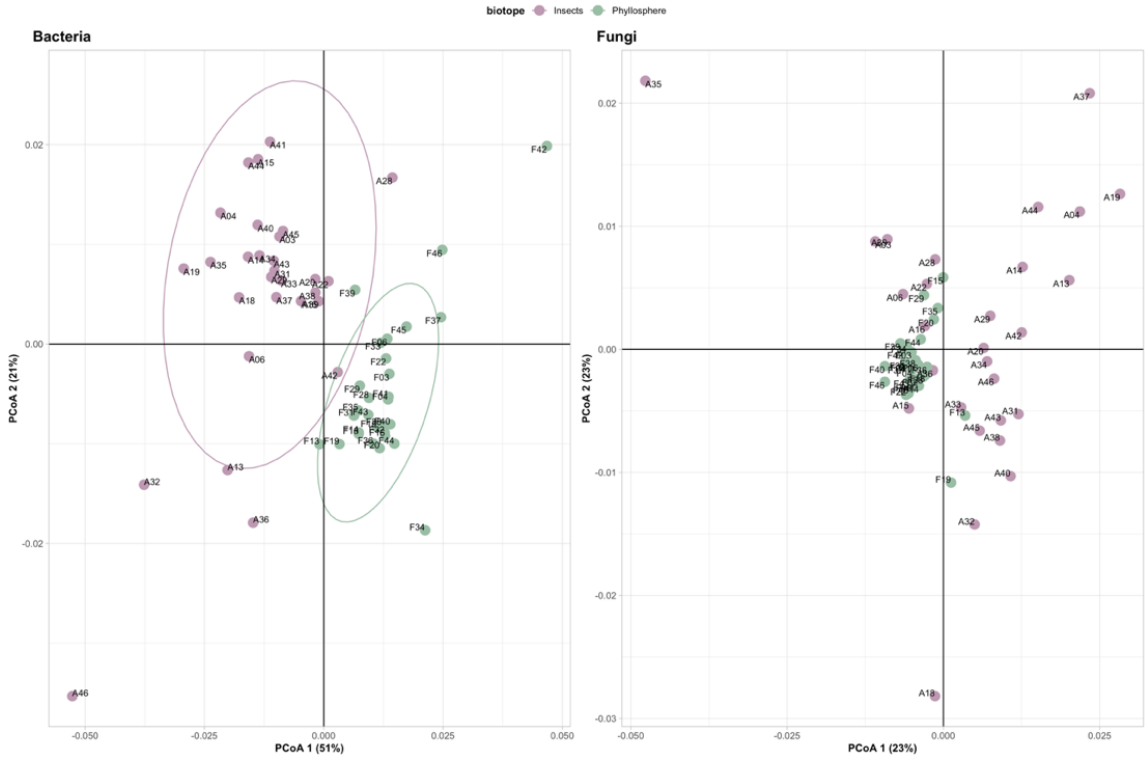


Figure 3.3 PCoA basée sur les distances MNTD utilisant les données de la communauté bactérienne, p-value < 0.001 (à gauche) et la communauté fongique, p-value < 0.0507 (à droite). Pour chaque biotope, le nombre d'échantillons considérés est de 29.

3.5.3 Les réseaux de cooccurrences des taxa dans les deux biotopes

Des 61 ASV impliqués dans l'analyse des cooccurrences pour le tract intestinal de l'insecte, 1830 paires de combinaison ont été créées, conduisant à 28,8% de cooccurrences positives (527 combinaisons), moins de 1% de cooccurrences négatives (12 combinaisons) et 70% de cooccurrences aléatoires (**Figure 3.4**). Deux ASV se sont révélés comme étant des nœuds clés pour ce réseau écologique avec des valeurs de degré de connexion et de proximité plus élevées (B189 et B31).

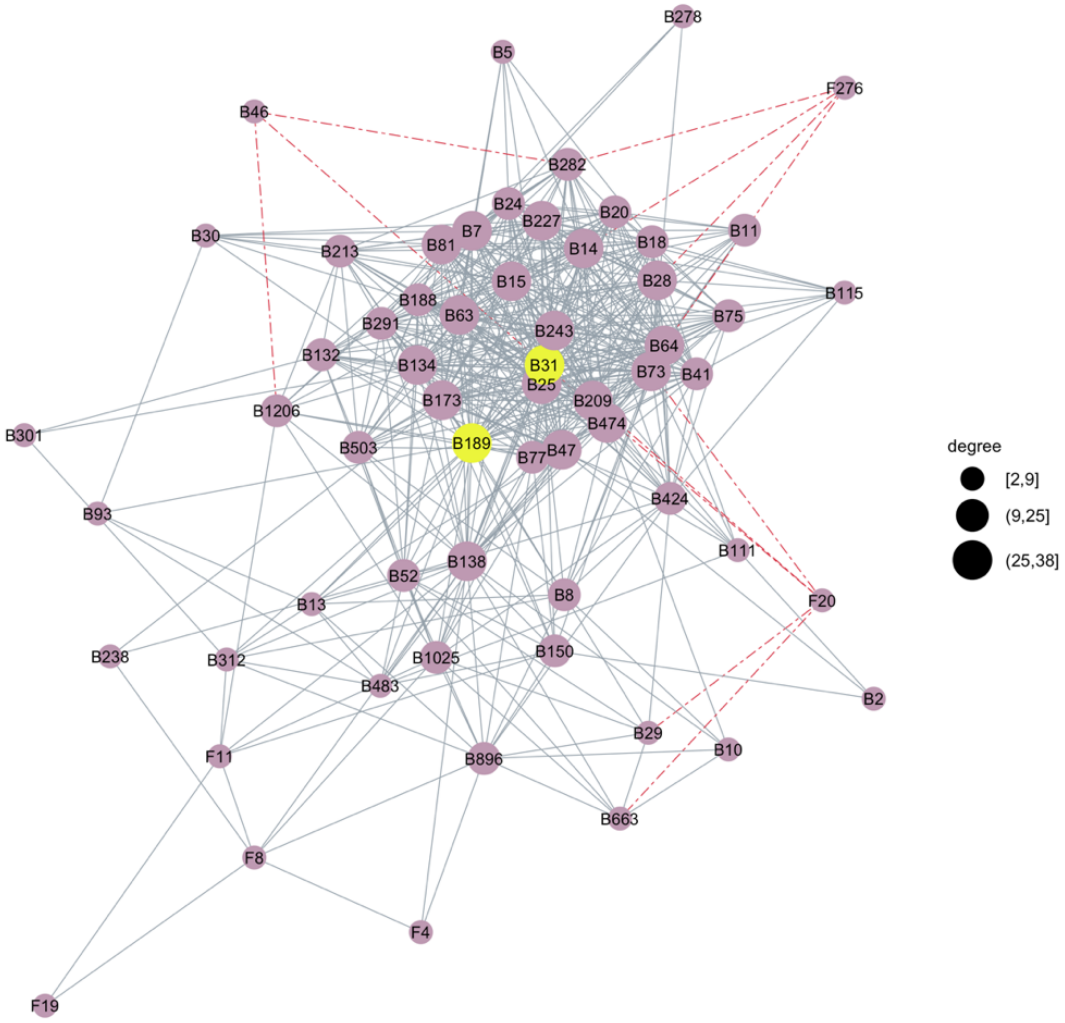


Figure 3.4 **Représentation des cooccurrences significatives dans le trac intestinal de l'insecte. Les nœuds en jaunes sont ceux ayant obtenu les valeurs de degré de connexion et de proximité les plus élevés (B31 et B189) correspondant à des taxa potentiellement clés pour le réseau de cooccurrences. En gris les cooccurrences significativement positives et en rouge les cooccurrences significativement négatives. La taille de chaque nœud (degré de connexion) est fonction du nombre de connexions arrivant à ce nœud.**

Dans le cas de la phyllosphère, à partir des 129 ASV impliqués dans l'analyse, les résultats montrent que 8256 paires de combinaisons, dont 92,9% étaient aléatoires, 6% positives (494 combinaisons) et 1% négatives (**Figure 3.5**). Comparativement au tract intestinal de l'insecte, deux ASV (B31320 et B885) potentiellement importants pour la stabilité de la structure écologie (valeurs de degré de connexion et de proximité les plus élevées).

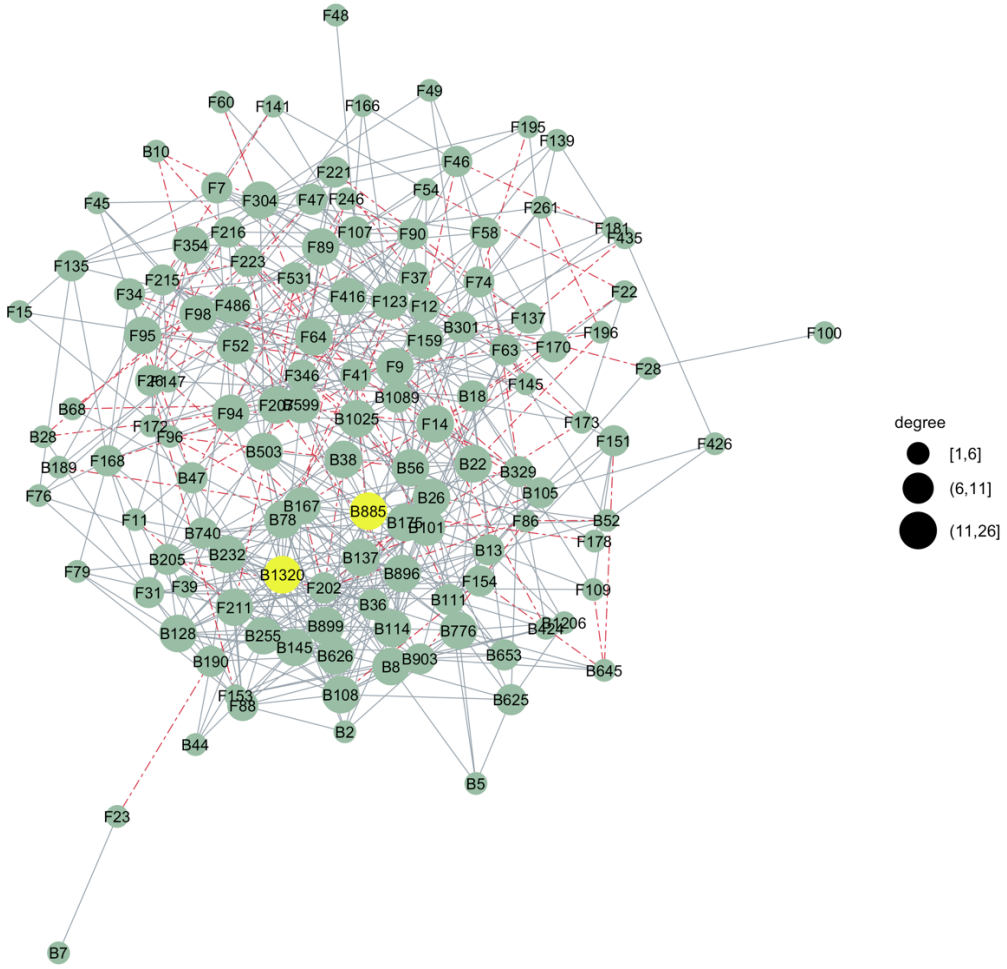


Figure 3.5 **Représentation des cooccurrences significatives dans la phyllosphère. Les nœuds en jaunes sont ceux ayant obtenu les valeurs de degré de connexion et de proximité les plus élevés (B31320 et B885) correspondant à des taxa potentiellement clés pour le réseau de cooccurrences. En gris les cooccurrences significativement positives et en rouge les cooccurrences significativement négatives. La taille de chaque nœud (degré de connexion) est fonction du nombre de connexions arrivant à ce nœud.**

Les deux métriques topologiques évaluées (degré de nœud et centralité de proximité) étaient plus élevées au niveau du tract intestinal de l'insecte indiquant un réseau écologique dans le tract intestinal plus dense que celui se trouvant sur la phyllosphère (**Figure 3.6**).

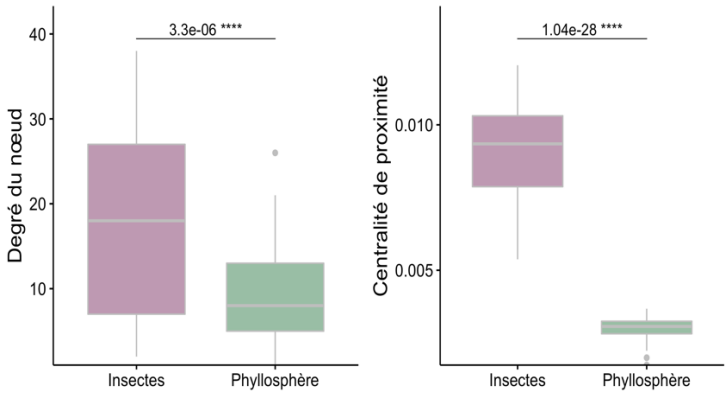
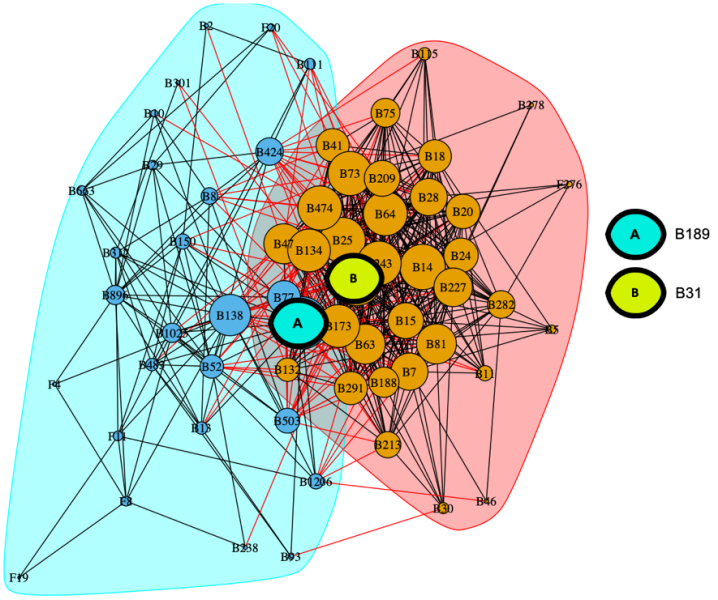


Figure 3.6 Comparaison des mesures de centralité chez les deux hôtes

L'analyse de la modularité a mis en évidence deux sous-réseaux, l'un mené par l'ASV B189 (Enterobacteriales, avec 27 ASV) et l'autre par l'ASV B31 (Bacillales, avec 34 ASV) dans le tract intestinal. Ce résultat vient conforter l'importance de ces deux ASV qui apparaissent comme des ASV concentrateurs (**Figure 3.7**). Le sous-réseau mené par l'ASV B31 (Enterobacteriales) était constitué majoritairement de Gammaprotéobactéries (59%), tandis que celui mené par l'ASV B189 (Bacillales) était constitué majoritairement d'Alphaprotéobactéries (33%). Au niveau de la phyllosphère, l'analyse de la modularité a révélé quatre sous-réseaux, menés chacun par des ASV concentrateurs : B1320 (Burkholderiales, avec 45 ASV), B105 (Pseudomonales, avec 15 ASV), F123 (Dothideomycetes, avec 44 ASV) et F9 (Pleosporales, avec 25 ASV). Les métriques calculées ont révélé que les quatre taxa avaient des valeurs les plus élevées dans chaque sous-réseau. Ainsi, l'ASV B1320 regroupait majoritairement des espèces bactériennes des classes de Bétaprotéobactéries (20%), Actinobactéries (20%) et Alphaprotéobactéries (20%) ; l'ASV B105 concentrait majoritairement des espèces des classes de Dothidéomycètes (47%) et Gammaprotéobactéries (27%) ; les ASV F123 et F9 concentraient majoritairement des espèces fongiques d'Ascomycètes avec respectivement 71% et 64%, essentiellement des Dothidéomycètes.

A- Tracts intestinaux des insectes (n=29)



B- Phyllosphère des arbres (n=29)

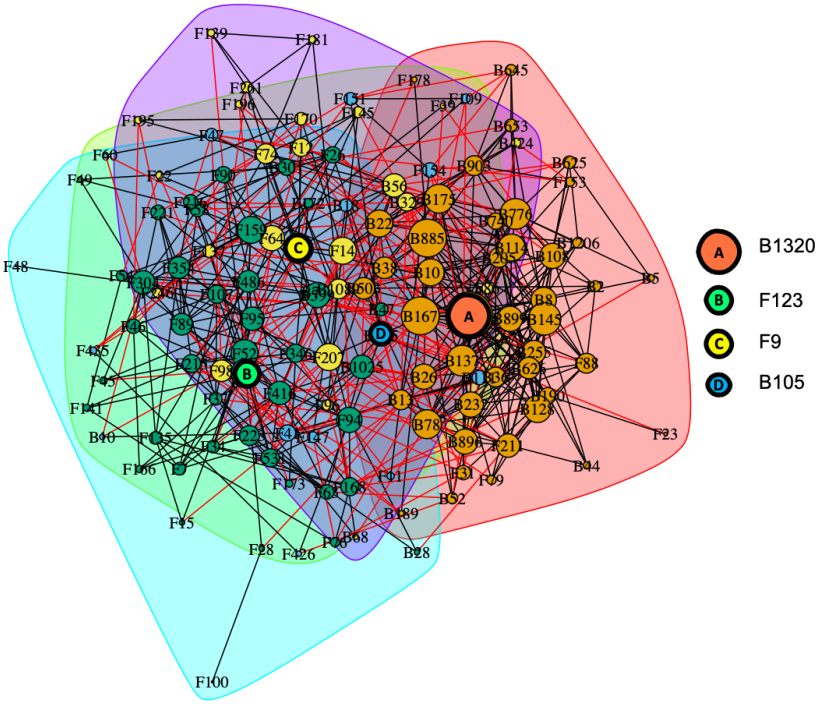


Figure 3.7 **Représentation de modularité dans les deux biotopes : A tracts intestinaux de l'insecte avec deux ASV concentrateurs (B189 et B31) ; B phyllosphère avec quatre ASV concentrateurs (B1320, B105, F123 et F9)**

3.6 Discussion et conclusion

L'analyse de la structure phylogénétique a été déterminée pour explorer les processus écologiques qui structurent les communautés microbiennes dans chaque biotope. Le regroupement phylogénétique peut être interprété comme un effet de sélection exercé par l'habitat avec des espèces étroitement apparentées, partageant un ou plusieurs traits leur permettant de persister dans cet environnement, favorisant une attraction phénotypique aux dépens de la répulsion (Webb, 2000; Webb *et al.*, 2002). Nos résultats suggèrent que les deux biotopes étudiés ont une importance déterminante sur le processus d'assemblage des communautés microbiennes avec une tendance au regroupement phylogénétique dans chaque environnement. La tendance au regroupement phylogénétique a récemment été documentée chez certains eucaryotes, notamment la communauté fongique du tract intestinal de *Achipteria coleoptrata*, ordre des Sarcoptiformes (Gong *et al.*, 2018), la communauté bactérienne associée au tract intestinal de certains mammifères herbivores (Youngblut *et al.*, 2019), la communauté bactérienne coexistant dans des habitats variés (mésocosme d'eau fraiche et sols de sources différentes) (Horner-Devine & Bohannan, 2006). Si la différenciation entre deux clades conduit à deux groupes de taxa qui ont des traits morphologiques assez différents entre eux. Ainsi, les taxa se trouvant dans le même groupe sont plus similaires à cause de leur ancêtre commun et de leur histoire évolutionnaire partagée (Soul & Wright, 2020). La différenciation quant à elle signifie que les taxa ayant subi la différenciation ont évolué séparément sur une période relativement longue et par conséquent ont pu accumuler des différences. De plus, la différenciation écologique est définie comme étant une différence aussi bien au niveau taxonomique qu'au niveau du contenu génétique (Hehemann *et al.*, 2016; Winans *et al.*, 2017). La prise en compte des données relatives à la structure phylogénétique dans l'analyse de la biodiversité dans les deux biotopes a permis d'apporter de la finesse dans nos résultats. En effet, l'article précédent avait permis de mettre en évidence une distinction significative entre les communautés microbiennes de chaque biotope, notamment avec une part d'explication attribuée à la structure phylogénétique dépassant les 70% pour la communauté bactérienne. Nos résultats indiquent que dans le tract intestinal de l'agrile du frêne adulte, les espèces bactériennes auraient connu cette différenciation écologique antérieure à celle des espèces fongiques. La tendance à faire moins de regroupement chez l'insecte (révélée par l'analyse du ratio NRI/NTI) a été confirmée par l'analyse des cooccurrences (analyse de la modularité). Très peu de données existent concernant la structure phylogénétique des microorganismes associés aux insectes. La littérature rapporte que la distinction entre les

espèces au sein du microbiome intestinal pourrait être attribuée d'une part aux conditions écologiques du tract intestinal, notamment à cause des conditions spécifiques de ce biotope (pH, présences de composés biologiques actifs, perturbations éventuelles) (Douglas, 2013), et d'autre part les interactions coévolutionnaires entre le microbiote et l'hôte pourraient entrainer une divergence évolutive (co-diversification) de ces microorganismes (Walter *et al.*, 2011). Il peut être intéressant d'explorer les relations potentielles entre la différenciation écologique de cette communauté bactérienne et les processus écologiques d'adaptation de l'insecte au moment de son établissement dans un nouvel environnement. Bien que ces résultats apportent des informations intéressantes, une exploration des corrélations potentielles entre ce phénomène de différenciation phylogénétique à des traits fonctionnels de l'insecte permettrait d'avoir une compréhension plus complète de l'implication des communautés microbiennes de l'insecte dans les processus d'adaptation écologiques chez celui-ci.

L'analyse probabiliste de cooccurrences a pu mettre en évidence les ASV significativement associés avec le tract intestinal de chaque hôte avec certaines connexions identifiées comme positives, d'autres, considérées négatives, et finalement un certain nombre d'interactions aléatoires. L'analyse des réseaux écologiques est souvent confrontée à un défi non négligeable qui est celui de l'interprétation des résultats. En effet, les cooccurrences identifiées pourraient être interprétées comme étant liées à de réelles interactions biologiques ou liées aux préférences des ASV pour l'hôte. D'après les travaux de Freilich *et al.* (2018), les réseaux de cooccurrences seraient être particulièrement pertinents lorsqu'on analyse la dynamique communautaire associant les interactions et l'environnement plutôt que les interactions par paires. Ainsi, les tentatives d'interprétation de nos résultats vont assumer que les cooccurrences identifiées sont potentiellement liées aux contraintes de l'hôte. La représentation des cooccurrences d'espèces chez leurs hôtes respectifs suggère un réseau beaucoup plus dense avec une connectivité plus élevée dans le tractus intestinal de l'insecte. Nos résultats suggèrent que les conditions imposées par l'habitat peuvent être déterminantes pour la distribution des taxa, notamment ici l'ASV B189 (Bacillales) dont la distribution est apparue différente selon l'hôte (38 cooccurrences positives dans le tract intestinal contre 2 cooccurrences négatives et une négative au niveau de la phyllosphère). Les cooccurrences identifiées étaient majoritairement positives concernant le tract intestinal contrairement aux observations au niveau de la phyllosphère de l'arbre hôte. Le tract intestinal a révélé une probabilité de cooccurrences significatives est plus élevée que dans la phyllosphère avec un réseau écologique plus dense (degré de nœud plus élevé). De plus certaines espèces sont apparues avec des valeurs de centralité particulièrement élevées,

suggérant que ces espèces pourraient être considérées comme des ASV clés dans les réseaux construits. En effet, d'après plusieurs travaux (Banerjee *et al.*, 2016; González *et al.*, 2010; Williams *et al.*, 2014), les nœuds (ASV) ayant des valeurs élevées des métriques explorées dans cette étude étaient importants dans le maintien de la connectivité des réseaux écologiques et pouvaient être considérés comme des espèces clés. Ces ASV sont apparus comme étant des éléments concentrateurs avec une tendance à former des sous-réseaux en fonction de son hôte. Certains travaux de suggèrent que ces ASV clés pourraient jouer un rôle important dans la structuration des communautés microbiennes, car ils pourraient entrainer la disparition ou le développement de certaines espèces (Agler *et al.*, 2016; Faust & Raes, 2012; van der Heijden & Hartmann, 2016). La modularité qui s'est révélée plus élevée pour les microorganismes de la phyllosphère était cohérente avec la densité plus élevée du réseau identifié dans le tract intestinal de l'insecte (degré de nœud plus élevé). (Montoya *et al.*, 2015) ont rapporté dans leurs travaux que la modularité était associée à la résistance aux perturbations. Ces mêmes travaux ont suggéré que la modularité des réseaux d'OTU cooccurrents pourrait être corrélée à la diversité en espèces, à la fonctionnalité et à leurs préférences environnementales. Très peu d'informations sont disponibles concernant l'assemblage des communautés bactériennes et fongiques associées au tract intestinal des insectes. Il pourrait être intéressant de conduire des analyses de corrélations entre les ASV clés identifiés et des paramètres tels que la taille de la population d'insectes chez l'arbre hôte, certains traits fonctionnels de l'insecte ayant des effets potentiels sur la structure et la fonction. Un exemple de corrélation a été rapporté dans les travaux de (Greenblum *et al.*, 2012) entre une modularité réduite et le microbiome des humains atteints d'obésité et de maladies intestinales inflammatoires, à travers des réseaux métaboliques. L'étude récente de Youngblut *et al.* (2019) a montré que la pression sélective de l'environnement et les interactions microbes-microbes différaient en fonction de l'hôte.

Tableau Suppl. 3.1: Résultats détaillés des valeurs de SES calculées par site. Les valeurs avec des astérisques (*) indiquent les sites ayant une structure phylogénétique de regroupement tandis que celles en gras et astérisque indiquent une structure phylogénétique de surdispersion significative.

	Bactéries			*Champignons*	
	Insectes NRI/NTI	Phyllosphère NRI/NTI		Insectes NRI/NTI	Phyllosphère NRI/NTI
Site 03	1,10	0,03	Site 03	1,91	0,49
Site 04	-0,03	0,48	Site 04	-2,73	0,75
Site 06	1,19	0,51	Site 06	0,63	0,69
Site 13	0,20	-0,08	Site 13	-0,03	0,72
Site 14	0,11	-0,09	Site 14	-0,26	0,71
Site 15	0,42	0,37	Site 15	-0,01	-3,17
Site 16	0,65	-0,05	Site 16	1,52	0,91
Site 18	0,40	0,02	Site 18	0,97	0,64
Site 19	0,13	0,09	Site 19	-0,49	0,67
Site 20	-1,38	-0,01	Site 20	0,90	0,45
Site 22	0,17	0,34	Site 22	1,35	0,56
Site 28	2,32	0,04	Site 28	-89,02	0,71
Site 29	0,12	0,46	Site 29	0,60	-0,48
Site 31	-0,47	-0,12	Site 31	0,22	0,64
Site 32	0,14	-0,33	Site 32	0,60	0,72
Site 33	0,44	0,28	Site 33	2,69	0,61
Site 34	-0,09	-0,17	Site 34	1,01	0,31
Site 35	0,40	0,77	Site 35	-0,29	0,90
Site 36	0,34	-0,07	Site 36	1,01	0,92
Site 37	0,94	0,05	Site 37	-0,04	0,49
Site 38	0,56	0,10	Site 38	2,68	-0,15
Site 39	0,16	-0,24	Site 39	0,19	0,63
Site 40	0,23	0,40	Site 40	0,49	-1,62
Site 41	0,33	-0,04	Site 41	0,14	0,53
Site 42	0,20	1,80	Site 42	1,15	0,48
Site 43	0,21	-0,46	Site 43	0,91	0,52
Site 44	-0,06	0,12	Site 44	1,06	1,16
Site 45	0,24	0,28	Site 45	1,28	0,82
Site 46	-0,01	0,17	Site 46	0,51	-2,68

BIBLIOGRAPHIE (ARTICLE 3)

Agler MT, Ruhe J, Kroll S, Morhenn C, Kim S-T, Weigel D, Kemen EM (2016) Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS biology* 14(1):e1002352.

Banerjee S, Kirkby CA, Schmutter D, Bissett A, Kirkegaard JA, Richardson AE (2016) Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. *Soil Biology and Biochemistry* 97:188-198.

Briatte F (2021) ggnet: Functions to plot networks with ggplot2.

Cadotte MW, Jonathan Davies T, Regetz J, Kembel SW, Cleland E, Oakley TH (2010) Phylogenetic diversity metrics for ecological communities: integrating species richness, abundance and evolutionary history. *Ecology letters* 13(1):96-105.

Carson W & Schnitzer S (2011) *Tropical forest community ecology.* John Wiley & Sons,

Clauset A, Newman ME, Moore C (2004) Finding community structure in very large networks. *Physical review E* 70(6):066111.

Csardi G & Nepusz T (2006) The igraph software package for complex network research. *InterJournal, complex systems* 1695(5):1-9.

Dillon R & Dillon V (2004) The gut bacteria of insects: nonpathogenic interactions. *Annual Reviews in Entomology* 49(1):71-92.

Douglas AE (2013) Microbial brokers of insect-plant interactions revisited. *Journal of chemical ecology* 39(7):952-961.

Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biological conservation* 61(1):1-10.

Faust K & Raes J (2012) Microbial interactions: from networks to models. *Nature Reviews Microbiology* 10(8):538-550.

Freeman LC, Roeder D, Mulholland RR (1979) Centrality in social networks: II. Experimental results. *Social networks* 2(2):119-141.

Freilich MA, Wieters E, Broitman BR, Marquet PA, Navarrete SA (2018) Species co-occurrence networks: Can they reveal trophic and non-trophic interactions in ecological communities? *Ecology* 99(3):690-699.

Gong X, Chen T-W, Zieger SL, Bluhm C, Heidemann K, Schaefer I, Maraun M, Liu M, Scheu S (2018) Phylogenetic and trophic determinants of gut microbiota in soil oribatid mites. *Soil Biology and Biochemistry* 123:155-164.

González AMM, Dalsgaard B, Olesen JM (2010) Centrality measures and the importance of generalist species in pollination networks. *Ecological complexity* 7(1):36-43.

Greenblum S, Turnbaugh PJ, Borenstein E (2012) Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proceedings of the National Academy of Sciences* 109(2):594-599.

Griffith DM, Veech JA, Marsh CJ (2016) Cooccur: probabilistic species co-occurrence analysis in R. *Journal of Statistical Software* 69(2):1-17.

Hehemann J-H, Arevalo P, Datta MS, Yu X, Corzett CH, Henschel A, Preheim SP, Timberlake S, Alm EJ, Polz MF (2016) Adaptive radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes. *Nature communications* 7(1):1-10.

Horner-Devine MC & Bohannan BJ (2006) Phylogenetic clustering and overdispersion in bacterial communities. *Ecology* 87(sp7):S100-S108.

Hunter PJ, Hand P, Pink D, Whipps JM, Bending GD (2010) Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (Lactuca species) phyllosphere. *Applied and environmental microbiology* 76(24):8117-8125.

Itoh H, Jang S, Takeshita K, Ohbayashi T, Ohnishi N, Meng X-Y, Mitani Y, Kikuchi Y (2019) Host–symbiont specificity determined by microbe–microbe competition in an insect gut. *Proceedings of the National Academy of Sciences* 116(45):22673-22682.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26(11):1463-1464.

Kembel SW & Hubbell SP (2006) The phylogenetic structure of a neotropical forest tree community. *Ecology* 87(sp7):S86-S99.

Layeghifard M, Hwang DM, Guttman DS (2017) Disentangling interactions in the microbiome: a network perspective. *Trends in microbiology* 25(3):217-228.

Mogouong J, Constant P, Lavallée R, Guertin C (2020) Gut microbiome of the emerald ash borer, Agrilus planipennis Fairmaire, and its relationship with insect population density. *FEMS Microbiology Ecology* 10.1093/femsec/fiaa141.

Mogouong J, Constant P, Legendre P, Guertin C (2021) The phyllosphere microbiome of host trees contributes more than leaf phytochemicals to variation in the Agrilus planipennis Fairmaire gut microbiome structure. *Scientific Reports* 11(1):1-15.

Montoya D, Yallop M, Memmott J (2015) Functional group diversity increases with modularity in complex food webs. *Nature communications* 6(1):1-9.

Newman ME & Girvan M (2004) Finding and evaluating community structure in networks. *Physical review E* 69(2):026113.

Olesen JM, Bascompte J, Dupont YL, Jordano P (2007) The modularity of pollination networks. *Proceedings of the National Academy of Sciences* 104(50):19891-19896.

Simon J-C, Marchesi JR, Mougel C, Selosse M-A (2019) Host-microbiota interactions: from holobiont theory to analysis. *Microbiome* 7(1):1-5.

Soul L & Wright D (2020) Phylogenetic Comparative Methods: A User's Guide for Paleontologists.

Vamosi JC & Vamosi SM (2007) Body size, rarity, and phylogenetic community structure: insights from diving beetle assemblages of Alberta. *Diversity and Distributions* 13(1):1-10.

van der Heijden MG & Hartmann M (2016) Networking in the plant microbiome. *PLoS Biology* 14(2):e1002378.

van Veelen HPJ, Salles JF, Tieleman BI (2017) Multi-level comparisons of cloacal, skin, feather and nest-associated microbiota suggest considerable influence of horizontal acquisition on the microbiota assembly of sympatric woodlarks and skylarks. *Microbiome* 5(1):1-17.

Veech JA (2013) A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography* 22(2):252-260.

Walter J, Britton RA, Roos S (2011) Host-microbial symbiosis in the vertebrate gastrointestinal tract and the Lactobacillus reuteri paradigm. *Proceedings of the National Academy of Sciences* 108(Supplement 1):4645-4652.

Webb CO (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist* 156(2):145-155.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ (2002) Phylogenies and community ecology. *Annual review of ecology and systematics* 33(1):475-505.

Williams RJ, Howe A, Hofmockel KS (2014) Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. *Frontiers in microbiology* 5:358.

Winans NJ, Walter A, Chouaia B, Chaston JM, Douglas AE, Newell PD (2017) A genomic investigation of ecological differentiation between free-living and Drosophila-associated bacteria. *Molecular ecology* 26(17):4536-4550.

Youngblut ND, Reischer GH, Walters W, Schuster N, Walzer C, Stalder G, Ley RE, Farnleitner AH (2019) Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. *Nature communications* 10(1):1-15.

Yun J-H, Roh SW, Whon TW, Jung M-J, Kim M-S, Park D-S, Yoon C, Nam Y-D, Kim Y-J, Choi J-H (2014) Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Applied and environmental microbiology* 80(17):5254-5264.

4 DISCUSSION – CONCLUSION – PERSPECTIVES

Les questions de recherche sur le microbiome ont beaucoup évolué pendant les dernières décennies donnant lieu à une documentation de plus en plus prometteuse sur le microbiome du tract intestinal des insectes, notamment les travaux en lien avec les rôles bénéfiques pour leur hôte. Il a été rapporté que la plupart des tracts intestinaux des insectes contiennent peu d'espèces microbiennes comparées au tract intestinal des mammifères (Engel & Moran, 2013). L'ensemble des travaux de cette thèse avaient pour objectif principal d'apporter des éléments de compréhension relatifs aux modulations du microbiome du tract intestinal des adultes de l'agrile du frêne, lorsqu'ils se retrouvent confrontés à des facteurs environnementaux. L'agrile du frêne qui a servi de modèle d'étude est très peu documenté. Cet insecte représente non seulement une réelle menace pour plusieurs espèces de frênes en Amérique du Nord dont certaines endémiques (Klooster *et al.*, 2018), mais aussi des espèces de frênes présents en Europe (Russie et Ukraine) (Orlova-Bienkowskaja *et al.*, 2020). Son établissement depuis son arrivée sur le continent n'a pas pu être maitrisé, et des travaux ont rapporté un succès d'établissement sur la base d'observations qui font état de comportements très variables en fonction de facteurs tels que la température (un exemple documenté est le cycle biologique dont la durée est variable selon la température). Le rôle important et parfois essentiel du microbiome pour son hôte présente un intérêt indéniable devrait être considéré. Les approches utilisées pour les travaux de thèse ont permis de mettre en évidence un aperçu des communautés bactérienne et fongique associées au tract intestinal de l'agrile du frêne adulte, tout en suggérant une certaine plasticité du microbiome intestinal de l'insecte confrontée au proxi de la densité de la population au sein des arbres hôtes. Ainsi, la variation de la composition microbienne dans l'intestin de l'agrile du frêne pourrait être liée à l'adaptation aux propriétés de l'arbre hôte.

Des communautés bactérienne et fongique relativement diversifiées : Plusieurs travaux ont permis de caractériser le microbiome du tract intestinal chez divers insectes en utilisant l'approche de séquençage à haut débit, mais c'est la première fois que ce travail est effectué sur l'agrile du frêne (modèle d'étude) en considérant simultanément deux des communautés microbiennes associées au tract intestinal de l'insecte (bactéries et champignons). La communauté bactérienne du tract intestinal de l'agrile adulte est apparue avec richesse spécifique plus élevée que celle de la communauté fongique. D'une part, la diversité alpha au sein de la communauté bactérienne s'est révélée sensible au proxi « niveau de densité des populations dans l'arbre hôte », d'autre part elle est apparue avec une différenciation écologique

plus précoce que la communauté fongique suggérant un rôle important de cette communauté bactérienne dans des fonctions spécialistes. À l'inverse la communauté fongique est apparue plus hétérogène, avec une majorité des taxa identifiés aussi présents au niveau de la phyllosphère. L'approche utilisée pour cette étude n'a malheureusement pas permis différencier de manière certaine le microbiome indigène propre à l'insecte du microbiome transitoire qui transite dans le tract intestinal notamment par les matières ingérées. En ce sens, des études supplémentaires seront nécessaires pour confirmer cette tendance, notamment en utilisant par exemple une approche basée sur le séquençage à haut débit de l'ARN (RNA-seq) qui permet l'étude de l'ensemble des ARN (transcrits), généralement des ARN messagers et qui sont utilisés comme intermédiaires pour la production de protéines, le tout couplé à une méthode de PCR quantitative.

Le tract intestinal de l'insecte, une pression déterminante pour la sélection les microorganismes associés : Quelques travaux antérieurs ont rapporté l'influence de l'écologie de l'hôte sur la dynamique de son microbiote (Engel *et al.*, 2016; Santos-Garcia *et al.*, 2020). Nos résultats ont mis en évidence la plasticité du microbiome du tract intestinal de l'agrile du frêne. Ainsi, l'effet déterminant de deux facteurs sur l'assemblage des espèces composant le microbiome du tract intestinal de l'insecte, d'une part l'environnement propre de l'hôte en intégrant la dimension phylogénétique du microbiote et d'autre part les microorganismes associés à la source alimentaire de l'insecte. On pourrait ainsi penser que les facteurs influençant les traits physiologiques ou fonctionnels de l'insecte pourraient aussi indirectement influencer la structure des communautés microbiennes du tract intestinal de l'insecte. L'effet significatif de la taille des populations d'insectes au sein de l'arbre hôte sur la structure du microbiome intestinal pourrait être associé d'une part à des changements de traits fonctionnels ou physiologiques au niveau de l'insecte et d'autre part à des changements significatifs au niveau des propriétés de la phyllosphère de l'arbre hôte. Ces deux phénomènes peuvent être amorcés suite au déclenchement du mécanisme d'adaptation des insectes envahisseurs, notamment les insectes exotiques qui se retrouvent contraints de s'adapter à de nouvelles conditions environnementales. Plusieurs travaux ayant rapporté que les transferts horizontaux de certains microorganismes (symbiontes) pourraient servir de mécanisme puissant et immédiat d'adaptation rapide, se traduisant par une acquisition instantanée par les espèces exotiques de nouveaux traits écologiques importants, notamment la tolérance à la chaleur (Henry *et al.*, 2013; Moran *et al.*, 2008). De plus, Bohan *et al.* (2017) rapportent aussi dans leurs travaux que les insectes herbivores bénéficient régulièrement des associations avec des symbiontes facultatifs, à l'instar

des aphidés confrontés à des stress écologiques incluant des changements chez les plantes hôtes.

La pression du microbiome associé à la source alimentaire sur le microbiome de l'insecte pourrait-elle être déterminante ? Les travaux de cette thèse ont également suggéré une influence significative du microbiome provenant de la phyllosphère sur celui du tract intestinal de l'insecte avec des espèces cooccurrentes chez les deux hôtes. L'agrile du frêne est surtout réputé en Amérique du Nord par le succès de son établissement dans un nouvel environnement si différent de sa région d'origine et sa stratégie d'attaque ne permettant pas d'identifier les débuts d'infestation. Une question pertinente qui pourrait se poser est de savoir si des facteurs environnementaux tels que les microorganismes de la phyllosphère de l'arbre hôte peuvent contribuer dans la facilitation du processus d'adaptation de l'insecte. La difficulté de détection des arbres nouvellement infestés a été largement documentée (Anulewicz, 2006; McCullough, 2020), de même qu'il a été rapporté qu'avec l'apparition visible de signes d'infestation les frênes sont condamnés à mourir dans les 4-5 ans à venir (Klooster *et al.*, 2014; McCullough *et al.*, 2019). Il peut donc être intéressant d'explorer la piste relative à ces espèces cooccurrentes, notamment recherche s'il existe une corrélation entre la présence de ces espèces dans le tract intestinal et des changements de traits physiologiques ou fonctionnels chez l'insecte. Nos résultats ont aussi suggéré que d'autres propriétés propres aux arbres hôtes auraient un effet significatif sur la structure du microbiome intestinal de l'insecte, notamment la position géographique et la teneur en cellulose des feuilles. Bien que nécessitant des études complémentaires pour confirmer les tendances trouvées, ces résultats intéressants amènent à se poser des questions pertinentes relatives à l'implication potentielle du microbiome de l'insecte dans le processus d'établissement de l'insecte en Amérique du Nord. À ce jour, très peu d'études ont pu mettre en évidence la sensibilité du tract intestinal des insectes vis-à-vis de ces facteurs relatifs aux arbres hôtes. Les connaissances apportées par les travaux de cette thèse représentent une contribution significative qui devrait être prise en considération dans les futures approches de lutte contre les insectes ravageurs notamment les approches visant à la déstabilisation du microbiome intestinal de ces insectes.

Des relations microbes-microbes complexes dans le tract intestinal de l'insecte : Le réseau de cooccurrence identifié dans le tract intestinal de l'insecte est apparu très dense avec des taxa qui semblent jouer un rôle plus important que les autres. Ces résultats représentent un élément supplémentaire qui vient confirmer la complexité des relations microbes-microbes au sein du tract

intestinal qui pourrait également contribuer aux variations du microbiome. Des recherches futures sont nécessaires pour élucider le mécanisme de cette variation et, plus important encore, les conséquences de cette variabilité sur l'adaptation des insectes aux changements environnementaux. Il pourrait par exemple être intéressant d'explorer un déséquilibre potentiel du microbiome en élevant en laboratoire des insectes. Maintenant qu'un lien significatif a été établi entre des propriétés des arbres hôtes et la structure du microbiome intestinal de l'insecte, il peut être intéressant d'axer les études futures sur la relation entre le niveau d'infestation des arbres et certaines de leurs propriétés comme la structure du microbiome de la phyllosphère ou la teneur en cellulose des feuilles, ou encore sur le comportement des espèces jouant un rôle concentrateur et des traits physiologiques de l'insecte. En effet, des mises en situation en laboratoire pourraient permettre d'établir le lien causal mis en évidence dans les trois articles afin de confirmer que nos résultats correspondent bien à des réponses réelles face aux facteurs déterminants identifiés. De plus, de futures études seront nécessaires pour tenter d'établir si des caractères fonctionnels bénéfiques peuvent être fournis à l'agrile du frêne par l'ingestion de microbiotes associés aux feuilles de frêne. Ceci pourrait fournir de précieuses indications sur la contribution des microorganismes environnementaux à la condition physique des insectes et sur la façon dont cela pourrait être manipulé à l'avantage de l'homme dans le cas d'espèces nuisibles comme l'agrile du frêne.

Les différentes approches utilisées dans les travaux de cette thèse sont originales et novatrices dans le domaine de l'analyse des données de séquençage destinées à élucider des questions de recherche sur le microbiome. En effet, la majorité des travaux effectués jusqu'à ce jour visant à répondre à des questions relatives à la dynamique ou la plasticité du microbiome ne prenaient en considération qu'une communauté microbienne à la fois (bactérienne, fongique ou encore virale). Il est possible que certains travaux aient pu surestimer certains effets sur la communauté bactérienne par exemple. Nos résultats mettent en évidence la nécessité de privilégier une approche globale incluant plusieurs paramètres à travers une analyse de type « partitionnement de la variation ». Cette originalité permet de donner un peu plus de profondeur dans l'interprétation et surtout d'avoir une meilleure compréhension de la complexité des processus mécaniques du microbiome des insectes.

BIBLIOGRAPHIE (CONCLUSION)

Anulewicz AC (2006) Field surveys of emerald ash borer (Agrilus planipennis Fairmaire)(Coleoptera: Buprestidae) density and canopy dieback in three North American ash species. *The 2006 ESA Annual Meeting, December 10-13, 2006.*

Bohan D, Dumbrell A, Massol F (2017) *Networks of Invasion: Empirical Evidence and Case Studies.* Academic Press,

Engel P, Kwong WK, McFrederick Q, Anderson KE, Barribeau SM, Chandler JA, Cornman RS, Dainat J, De Miranda JR, Doublet V (2016) The bee microbiome: impact on bee health and model for evolution and ecology of host-microbe interactions. *MBio* 7(2).

Engel P & Moran NA (2013) The gut microbiota of insects–diversity in structure and function. *FEMS microbiology reviews* 37(5):699-735.

Henry LM, Peccoud J, Simon J-C, Hadfield JD, Maiden MJ, Ferrari J, Godfray HCJ (2013) Horizontally transmitted symbionts and host colonization of ecological niches. *Current biology* 23(17):1713-1717.

Klooster WS, Gandhi KJ, Long LC, Perry KI, Rice KB, Herms DA (2018) Ecological impacts of emerald ash borer in forests at the epicenter of the invasion in North America. *Forests* 9(5):250.

Klooster WS, Herms DA, Knight KS, Herms CP, McCullough DG, Smith A, Gandhi KJ, Cardina J (2014) Ash (Fraxinus spp.) mortality, regeneration, and seed bank dynamics in mixed hardwood forests following invasion by emerald ash borer (Agrilus planipennis). *Biological Invasions* 16(4):859-873.

McCullough DG (2020) Challenges, tactics and integrated management of emerald ash borer in North America. *Forestry: An International Journal of Forest Research* 93(2):197-211.

McCullough DG, Poland TM, Tluczek AR, Anulewicz A, Wieferich J, Siegert NW (2019) Emerald ash borer (Coleoptera: Buprestidae) densities over a 6-yr period on untreated trees and trees treated with systemic insecticides at 1-, 2-, and 3-yr intervals in a Central Michigan Forest. *Journal of Economic Entomology* 112(1):201-212.

Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. *Annual review of genetics* 42:165-190.

Orlova-Bienkowskaja MJ, Drogvalenko AN, Zabaluev IA, Sazhnev AS, Peregudova EY, Mazurov SG, Komarov EV, Struchaev VV, Martynov VV, Nikulina TV (2020) Current range of Agrilus planipennis Fairmaire, an alien pest of ash trees, in European Russia and Ukraine. *Annals of Forest Science* 77(2):1-14.

Santos-Garcia D, Mestre-Rincon N, Zchori-Fein E, Morin S (2020) Inside out: microbiota dynamics during host-plant adaptation of whiteflies. *The ISME Journal* 14(3):847-856.

5 ANNEXES

5.1 Annexe 1 : Pipeline pour le traitement des données de séquençage

Adaptation du pipeline proposé par Edgar (2010)
(https://www.drive5.com/usearch/manual/pipe_examples.html)

Étape 1 : merging des paired-end reads
La commande –fastq_mergpairs a été utilisé pour assembler les séquences et créer des séquences consensus. Cette commande a plusieurs options à ajuster en fonction du type de data à analyse (origine des séquences). *–fast maxdiffs* correspond au nombre maximum de mismatches dans l'alignement. Par défaut il est à 5, il est recommandé de l'augmenter si les overlaps sont longs. **./usearch -fastq_mergepairs ./data/*_R1.fastq -relabel @ -fastq_maxdiffs 5 -fastqout raw.fq**

Étape 2 : Retrait des amorces
Permet de couper les amorces sur les deux côtés (forward et reverse). Il faut rajouter le nombre de nucléotides des amorces juste après la commande *–stripleft* et *–stripright*. Dans notre cas, les amorces utilisées pour les champignons avaient 18 bases.
./usearch -fastx_truncate raw.fq -stripleft 18 -stripright 18 -fastqout stripped.fq

Étape 3 : Filtration qualité
Permet de filtrer par les erreurs attendues (espérées : Phred scores (2 à 40), Expected number of errors (maxee), Overlapping paired reads) pour toutes les bases pendant la lecture et après les troncatures.
./usearch -fastq_filter stripped.fq -fastq_maxee 1.0 -fastaout filtered.fa -relabel Filt

Étape 4 : Déreplication
Sert à trouver toutes les séquences *uniques* dans le fichier d'entrée (fasta/fastq). Les séquences sont comparées lettre par lettre et doivent être identiques sur toute la longueur des deux séquences. La commande *–minuniquesize* établit un minimum d'abondance et les séquences *uniques* avec une abondance plus faible que le minimum, sont écartés. Par défaut, ce minimum est de 1. **./usearch -fastx_uniques filtered.fa -sizeout -minuniquesize 2 -relabel Uniq -fastaout uniques.fa**

Étape 5 : Clustering des otus
La commande *–cluster_otu* permet de faire un clustering en regroupant par OTU les séquences présantant au moins 97% de similitude en utilisant un algorythme. Les chimères sont aussi filtrées avec cette commande. La commande *–minsize* peut être utilisée pour spécifier l'abondance minimale. La valeur par défaut est de 2 (car élimine les singletons)
./usearch -cluster_otus uniques.fa -otus otus.fa -relabel Otu

Étape 6 : Création de la table otus
La commande *–otutab* permet de générer une table d'OTU en mappant les reads d'OTUs. **./usearch -otutab raw.fq -otus otus.fa -otutabout otutab_raw.txt**

Étape 7 : Normalisation de la table otus

Normalisation de la table d'OTUs à un même nombre de reads par échantillon. Elle peut être faite de plusieurs façons. Pour ce pipeline utilisé, la plus faible valeur de la taille de l'échantillon a été utilisée pour la normalisation. Pour avoir cette valeur, un rapport a été généré à cette étape, avant de lancer la commande de normalisation. Les comptes avec une fréquence plus faible que 0,5% du compte total ont été enlevés (Bokulich *et al.*, 2013). Les commandes utilisées sont *–otutab_trim* et *–min_otu_freq*.

./usearch -otutab_trim otutab_raw.txt -min_otu_freq 0.00005 -output trimmed.txt

Étape 8 : Génération d'un rapport pour la suite

Ce rapport permet de fixer les paramètres de la standardisation.

./usearch -otutab_stats trimmed.txt -output report.txt

Étape 9 : Standardisation de la table otus

Standardisation des comptes de la table d'OTUs au même nombre de reads par échantillon. La valeur de référence doit tenir compte des informations fournis dans le rapport de l'étape 8.

usearch -otutab_norm trimmed.txt -sample_size xxxxxx -output otutab.txt

Étape 10 : génération des indices de diversité

./usearch -alpha_div otutab.txt -output alpha.txt
./usearch -alpha_div_rare otutab.txt -output alpha_rare.txt

Étape 11 : génération de la table des taxas avec un cut-off établi à 80% de similitude

./usearch -sintax otus.fa -db ./data/rdp_its_v2.fa -strand both -tabbedout sintax.txt -sintax_cutoff 0.8 ./usearch -sintax_summary sintax.txt -otutabin otutab.txt -rank g -output genus_summary.txt
./usearch -sintax_summary sintax.txt -otutabin otutab.txt -rank p -output phylum_summary.txt

Références :

Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, Mills DA & Caporaso JG (2013) Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nature methods* 10(1):57-59.

5.2 Annexe 2 : Code R utilisé pour l'article # 1

```
##1- Import libraries
library(vegan)
library(ggplot2)
library(gridExtra)
library(plyr)
library(microbiome)
library(phyloseq)
library(clustsig)
library(ggdendro)
library(phyloseq)
library(stats)
library(ade4)
library(ape)
library(iNEXT)
library(adegenet)
library(phangorn)
library(adespatial)
library(indicspecies)
library(corrplot)
library(tidyverse)
library(dplyr)

##2- Diversity index
# diversity index of Bacteria
speb <- read.csv2("asv16_rawb.csv", row.names = 1)
outb <- iNEXT(t(speb), q = 0, datatype = "abundance", size = NULL, endpoint
= 25000, knots = 35, se = TRUE, conf = 0.95, nboot = 50)
ggiNEXT(outb, type = 1) # diversity
ggiNEXT(outb, type = 2) # coverage
Chao1_b = ChaoRichness(t(speb), datatype = "abundance", conf = 0.95)
summary(Chao1_b)
Simpson_b = ChaoSimpson(t(speb), datatype = "abundance", transform = F,
conf = 0.95)
summary(Simpson_b)
Shannon_b <- ChaoShannon(t(speb), datatype = "abundance", transform =
F,conf = 0.95)
summary(Shannon_b)

# diversity index of Fungi
spef <- read.csv2("asv16_rawf.csv", row.names = 1)
outf <- iNEXT(t(spef), q = 0, datatype = "abundance", size = NULL, endpoint
= 50000, knots = 35, se = TRUE, conf = 0.95, nboot = 50)
ggiNEXT(outf, type = 1) # diversity
ggiNEXT(outf, type = 2) # coverage
Chao1_f <- ChaoRichness(t(spef), datatype = "abundance", conf = 0.95)
summary(Chao1_f)
Simpson_f <- ChaoSimpson(t(spef), datatype = "abundance", transform = F,
conf = 0.95)
summary(Simpson_f)
Shannon_f <- ChaoShannon(t(spef), datatype = "abundance", transform =
F,conf = 0.95)
summary(Shannon_f)

# Chao1 Bacteria
div <- read.csv2("divb17.csv", row.names = 1)
summary_data <- ddply(div, c("Class"),summarise, N =
length(Chao1_Estimator), avg_value = mean(Chao1_Estimator), sd =
sd(Chao1_Estimator), se = sd/sqrt(N))
summary_data
```

```
raw_plotc <- ggplot(div, aes(x=reorder(Class,-
Chao1_Estimator),y=Chao1_Estimator))
raw_plotc <- raw_plotc + geom_point(position = position_jitter(w=0.1),
alpha=0.4, colour="grey58")
raw_plotc <- raw_plotc + xlab("Class") +  ylab("Chao1 Estimator (q=0)") +
theme_classic(base_size=18) + ylim(0,120)
combined_plotc <- raw_plotc + geom_point(data=summary_data, aes(x=Class,
y=avg_value), colour="black", size=3)
combined_plotc <- combined_plotc + geom_errorbar(data=summary_data, aes(x =
Class, y=avg_value,ymax=avg_value + se,ymin =avg_value - se), width=0.1,
colour="black")
combined_plotc <- combined_plotc + annotate("text",x="Low",y=0,label="n=3")
combined_plotc <- combined_plotc +
annotate("text",x="Intermediate",y=0,label="n=3")
combined_plotc <- combined_plotc +
annotate("text",x="High",y=0,label="n=3")
anova <- aov(Chao1_Estimator ~ Class, data=div)
summary(anova)
TukeyHSD(anova)
summary_data$labels <- c("b","ab","a")
summary_data$Class
combined_plotc <- combined_plotc + geom_point(data=summary_data,
aes(x=Class,y=avg_value), colour="black", size=3)
combined_plotc <- combined_plotc + geom_errorbar(data = summary_data, aes(x
= Class, y = avg_value, ymax = avg_value + se,ymin = avg_value - se),
width=0.1, colour="black")
combined_plot_chao <- combined_plotc +
geom_text(data=summary_data,aes(x=Class,y=avg_value+se,label=labels, vjust
= -4, hjust = 1))
```

 # Simpson Bacteria

```
summary_data <- ddply(div, c("Class"), summarise, N =
length(Simpson_Estimator), avg_value = mean(Simpson_Estimator), =
sd(Simpson_Estimator), se = sd/sqrt(29))
summary_data
raw_plotS <- ggplot(div, aes(x=Class,y=Simpson_Estimator)) # raw_data is
the name of the data you're using
raw_plotS <- raw_plotS + geom_point(position = position_jitter(w=0.1),
alpha=0.4, colour="grey58") # you are plotting points
raw_plotS <- raw_plotS + xlab("Class") +  ylab("Simpson Estimator (q=2)") +
theme_classic(base_size=18) + ylim(0,1) # label the axes
# add error bars
combined_plotS <- raw_plotS + geom_point(data=summary_data, aes(x=Class,
y=avg_value), colour="black", size=3)
combined_plotS <- combined_plotS + geom_errorbar(data=summary_data, aes(x =
Class, y=avg_value,ymax=avg_value + se,ymin =avg_value - se), width=0.1,
colour="black")
combined_plotS <- combined_plotS + annotate("text",x="Low",y=0,label="n=3")
combined_plotS <- combined_plotS +
annotate("text",x="Intermediate",y=0,label="n=3")
combined_plotS <- combined_plotS +
annotate("text",x="High",y=0,label="n=3")
anova <- aov(Simpson_Estimator ~ Class, data=div)
summary(anova)
TukeyHSD(anova)
summary_data$labels <- c("a","b","a")
summary_data$Class
```

```
combined_plotS <- combined_plotS + geom_point(data=summary_data,
aes(x=Class,y=avg_value), colour="black", size=3)
combined_plotS <- combined_plotS + geom_errorbar(data = summary_data, aes(x
= Class, y = avg_value, ymax = avg_value + se,ymin = avg_value - se),
width=0.1, colour="black")
combined_plot_simp <- combined_plotS +
geom_text(data=summary_data,aes(x=Class,y=avg_value+se,label=labels, vjust
= -2, hjust = 1))

# Shannon Bacteria
summary_data <- ddply(div, c("Class"), summarise, N =
length(Shannon_Estimator), avg_value = mean(Shannon_Estimator), sd =
sd(Shannon_Estimator), se = sd/sqrt(N))
summary_data
raw_plotN <- ggplot(div, aes(x=Class,y=Shannon_Estimator))
raw_plotN <- raw_plotN + geom_point(position = position_jitter(w=0.1),
alpha=0.4, colour="grey58")
raw_plotN <- raw_plotN + xlab("Class") +  ylab("Shannon Estimator (q=1)") +
theme_classic(base_size=18) + ylim(0,5)
combined_plotN <- raw_plotN + geom_point(data=summary_data, aes(x=Class,
y=avg_value), colour="black", size=3)
combined_plotN <- combined_plotN + geom_errorbar(data=summary_data, aes(x =
Class, y=avg_value,ymax=avg_value + se,ymin =avg_value - se), width=0.1,
colour="black")
combined_plotN <- combined_plotN + annotate("text",x="Low",y=0,label="n=3")
combined_plotN <- combined_plotN +
annotate("text",x="Intermediate",y=0,label="n=3")
combined_plotN <- combined_plotN +
annotate("text",x="High",y=0,label="n=3")
anova <- aov(Shannon_Estimator ~ Class, data=div)
summary(anova)
TukeyHSD(anova)
summary_data$labels <- c("a","a","a")
summary_data$Class
combined_plotN <- combined_plotN + geom_point(data=summary_data,
aes(x=Class,y=avg_value), colour="black", size=3)
combined_plot <- combined_plotN + geom_errorbar(data = summary_data, aes(x
= Class, y = avg_value, ymax = avg_value + se,ymin = avg_value - se),
width=0.1, colour="black")
combined_plot_shan <- combined_plotN +
geom_text(data=summary_data,aes(x=Class,y=avg_value+se,label=labels, vjust
= -2, hjust = 1))

#combine all the plots in the same figure
grid.arrange(combined_plot_chao, combined_plot_shan, combined_plot_simp,
ncol=3, nrow =1)

##3- Core microbiome
otumat <- read.csv2("asv16.csv", row.names = 1)
taxmat <- read.csv2("taxa16.csv", row.names = 1)
#rownames(otumat)
#rownames(taxmat)
otumat <- as.matrix(otumat)
taxmat <- as.matrix(taxmat)
#class(otumat)
#class(taxmat)
OTU.bact <- otu_table(t(otumat), taxa_are_rows = TRUE)
TAX.bact <- tax_table(taxmat)
dim(OTU.bact)
dim(TAX.bact)
```

```
bact <- phyloseq(OTU.bact, TAX.bact)
core.bact <- core(bact,0, 80/100)
core.tab <- otu_table(core.bact)
core.tax <- tax_table(core.bact)
dim(as.matrix(core.tab))
core.tab # ASV belonging to the core microbiome
core.tax # ASVs' taxa belonging to the core microbiome

##4- Hierarchical clustering UPGMA (Bray-Curtis distance)
spe <- read.csv2("asv16_trim1b.csv", row.names = 1)
spe.hel <- decostand(spe, method="hellinger")
spe.dhel <- vegdist(spe.hel,method="bray")
results <- simprof(spe.dhel, num.expected=1000, num.simulated=9999,
method.cluster="average",  method.transform="identity",
alpha=0.05,sample.orientation="row", const=0, silent=TRUE, increment=100,
undef.zero=TRUE,  warn.braycurtis=TRUE)
simprof.plot(results, leafcolors=c("red", "blue"), plot=TRUE, fill=TRUE,
leaflab="perpendicular", siglinetype=1)
results$significantclusters
summary(results)
results$numgroups

##5- Betadispersion - Permanova (Unifrac distance) - PCoA - Bacteria (ASV)
spe <- read.csv2("asv16_trim1b.csv", row.names = 1)
env <- read.csv2("env.csv", row.names = 1)

# betadispersion
spe.tr <- decostand(spe, "hel")
#dist.spe <- vegdist(spe.tr, method = "bray")
dist.spe = dist.ldc(spe.tr, method = "log.chord")
group <- as.factor(env$Insect_Density_Class)
bdisper <- betadisper(dist.spe, group, "centroid")
bdisper
anova(bdisper)
# for asv
#boxplot(bdisper, shape="group", main = "Homogeneity of groups dispersions
(bacteria)")

# Distance unifrac
otumat <- read.csv2("asv16_tree.csv", row.names = 1)
taxmat <- read.csv2("taxa16_tree.csv", row.names = 1)
rownames(otumat) <- c("FAD1b","FAD2b","FAD3b",
"MOD1b","MOD2b","MOD3b","FOD1b","FOD2b","FOD4b")
#rownames(otumat)
#rownames(taxmat)
otumat <- as.matrix(t(otumat))
taxmat <- as.matrix(taxmat)
OTU <- otu_table(otumat, taxa_are_rows = TRUE)
TAX <- tax_table(taxmat)
physeq <- phyloseq(OTU, TAX)
env <- read.csv2("env.csv", row.names = 1)
group <- as.factor(env$Insect_Density_Class)
EnvType <- sample_data(data.frame(group, row.names=sample_names(physeq),
stringsAsFactors=FALSE))
#EnvType
dna <- fasta2DNAbin(file="asvb.fa",quiet=TRUE)
D1 <- dist.dna(dna, model = "raw")
length(D1)
tre <- njs(D1)
class(tre)
physeq1 <- phyloseq(OTU, TAX, tre)
```

```
# permanova on unifrac distance (weighted)
distw <- distance(physeq1, "wunifrac", type = "samples")
PERMAw <- adonis2(distw ~ group, permutations = 9999, method = "unifrac")
PERMAw
spe<- as.matrix(spe)

# permanova on unifrac distance (unweighted)
distu <- distance(physeq1, "uunifrac", type = "samples")
PERMAu <- adonis2(distu ~ group, permutations = 9999, method = "unifrac")
PERMAu
spe <- as.matrix(spe)

# PCoA, weighted unifrac distance
spe <- t(otumat)
spe.pcoaw <- cmdscale(distw,  k=(nrow(spe)-1), eig=TRUE)
pc1w <- round(spe.pcoaw$eig[1]/sum(spe.pcoaw$eig)*100)
pc2w <- round(spe.pcoaw$eig[2]/sum(spe.pcoaw$eig)*100)
pc3w <- round(spe.pcoaw$eig[3]/sum(spe.pcoaw$eig)*100)
pc4w <- round(spe.pcoaw$eig[4]/sum(spe.pcoaw$eig)*100)
spe.pcoaw <- as.data.frame(spe.pcoaw$points)
spe.pcoaw <- spe.pcoaw[,1:3]
finalw <- cbind(spe.pcoaw,group)
PcoA_plot_asvbw <- ggplot(finalw, aes(x = V1, y = V2, colour = group, label
= row.names(spe))) + geom_point(size=9, show.legend = TRUE) + theme_light()
+ geom_hline(yintercept=0) + geom_vline(xintercept=0) +
theme(legend.position = "top")+ labs(x=paste("PCoA 1 (", pc1w, "%)", sep =
""), y=paste("PCoA 2 (", pc2w, "%)", sep = "")) + theme(panel.background =
element_blank(), plot.background=element_rect(fill="white",colour=NA)) +
geom_text(size = 1)+ ggtitle(label = "PCoA performed with bacteria
(weighted Unifrac)")
PcoA_plot_asvbw

# PCoA,unweighted unifrac distance
spe <- t(otumat)
spe.pcoau <- cmdscale(distu,  k=(nrow(spe)-1), eig=TRUE)
pc1u <- round(spe.pcoau$eig[1]/sum(spe.pcoau$eig)*100)
pc2u <- round(spe.pcoau$eig[2]/sum(spe.pcoau$eig)*100)
pc3u <- round(spe.pcoau$eig[3]/sum(spe.pcoau$eig)*100)
pc4u <- round(spe.pcoau$eig[4]/sum(spe.pcoau$eig)*100)
spe.pcoau <- as.data.frame(spe.pcoau$points)
spe.pcoau <- spe.pcoau[,1:3]
finalu <- cbind(spe.pcoau,group)
PcoA_plot_asvbu <- ggplot(finalu, aes(x = V1, y = V2, colour = group, label
= row.names(spe))) + geom_point(size=9, show.legend = TRUE) + theme_light()
+ geom_hline(yintercept=0) + geom_vline(xintercept=0) +
theme(legend.position = "top")+ labs(x=paste("PCoA 1 (", pc1u, "%)", sep =
""), y=paste("PCoA 2 (", pc2u, "%)", sep = "")) + theme(panel.background =
element_blank(), plot.background=element_rect(fill="white",colour=NA))+
geom_text(size = 1)+ ggtitle(label = "PCoA performed with bacteria
(unweighted Unifrac)")
PcoA_plot_asvbu

##5- Betadispersion - Permanova (Unifrac distance) - PCoA - Bacteria (OTU)
spe <- read.csv2("otu16b_tree.csv", row.names = 1)
env <- read.csv2("env.csv", row.names = 1)

# betadispersion
spe.tr <- decostand(spe, "hel")
#dist.spe <- vegdist(spe.tr, method = "bray")
dist.spe <- dist.ldc(spe.tr, method = "log.chord")
group <- as.factor(env$Insect_Density_Class)
```

```
bdisper <- betadisper(dist.spe, group, "centroid")
bdisper
anova(bdisper)
# for asv
# Distance unifrac
otumat <- read.csv2("otu16b_tree.csv", row.names = 1)
taxmat <- read.csv2("taxa16b_tree.csv", row.names = 1)
#otumat <- t(otumat)
rownames(otumat) = c("FAD1b","FAD2b","FAD3b",
"MOD1b","MOD2b","MOD3b","FOD1b","FOD2b","FOD4b")
rownames(otumat)
#rownames(taxmat)
otumat <- as.matrix(t(otumat))
taxmat <- as.matrix(taxmat)
OTU <- otu_table(otumat, taxa_are_rows = TRUE)
TAX <- tax_table(taxmat)
physeq <- phyloseq(OTU, TAX)
env <- read.csv2("env.csv", row.names = 1)
group <- as.factor(env$Insect_Density_Class)
EnvType <- sample_data(data.frame(group, row.names=sample_names(physeq),
stringsAsFactors=FALSE))
#EnvType
dna <- fasta2DNAbin(file="otub.fa",quiet=TRUE)
D1 <- dist.dna(dna, model = "raw") # TS
length(D1)
tre <- njs(D1)
class(tre)
physeq1 <- phyloseq(OTU, TAX, tre)

# permanova on weighted unifrac distance
distw <- distance(physeq1, "wunifrac", type = "samples")
PERMAw <- adonis2(distw ~ group,  permutations = 9999, method = "unifrac")
PERMAw
spe <- as.matrix(spe)

# permanova on unweighted unifrac distance
distu <- distance(physeq1, "uunifrac", type = "samples") # uunifrac
PERMAu <- adonis2(distu ~ group,  permutations = 9999, method = "unifrac")
PERMAu
spe <- as.matrix(spe)

# PCoA, weighted unifrac distance
spe <- t(otumat)
spe.pcoaw <- cmdscale(distw,  k=(nrow(spe)-1), eig=TRUE)
pc1w <- round(spe.pcoaw$eig[1]/sum(spe.pcoaw$eig)*100)
pc2w <- round(spe.pcoaw$eig[2]/sum(spe.pcoaw$eig)*100)
pc3w <- round(spe.pcoaw$eig[3]/sum(spe.pcoaw$eig)*100)
pc4w <- round(spe.pcoaw$eig[4]/sum(spe.pcoaw$eig)*100)
spe.pcoaw <- as.data.frame(spe.pcoaw$points)
spe.pcoaw <- spe.pcoaw[,1:3]
finalw <- cbind(spe.pcoaw,group)
PcoA_plot_asvbw <- ggplot(finalw, aes(x = V1, y = V2, colour = group, label
= row.names(spe))) + geom_point(size=9, show.legend = TRUE) + theme_light()
+ geom_hline(yintercept=0) + geom_vline(xintercept=0) +
theme(legend.position = "top")+ labs(x=paste("PCoA 1 (", pc1w, "%)", sep =
""), y=paste("PCoA 2 (", pc2w, "%)", sep = ""))+  theme(panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA))+
geom_text(size = 1)+ ggtitle(label = "PCoA performed with bacteria
(weighted Unifrac)")
PcoA_plot_asvbw
```

```
# PCoA,unweighted unifrac distance
spe <- t(otumat)
spe.pcoau <- cmdscale(distu,  k=(nrow(spe)-1), eig=TRUE)
pc1u <- round(spe.pcoau$eig[1]/sum(spe.pcoau$eig)*100)
pc2u <- round(spe.pcoau$eig[2]/sum(spe.pcoau$eig)*100)
pc3u <- round(spe.pcoau$eig[3]/sum(spe.pcoau$eig)*100)
pc4u <- round(spe.pcoau$eig[4]/sum(spe.pcoau$eig)*100)
spe.pcoau <- as.data.frame(spe.pcoau$points)
spe.pcoau <- spe.pcoau[,1:3]
finalu <- cbind(spe.pcoau,group)
PcoA_plot_asvbu <- ggplot(finalu, aes(x = V1, y = V2, colour = group, label
= row.names(spe))) + geom_point(size=9, show.legend = TRUE) + theme_light()
+ geom_hline(yintercept=0) + geom_vline(xintercept=0) +
theme(legend.position = "top")+ labs(x=paste("PCoA 1 (", pc1u, "%)", sep =
""), y=paste("PCoA 2 (", pc2u, "%)", sep = ""))+ theme(panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA))+
geom_text(size = 1)+ ggtitle(label = "PCoA performed with bacteria
(unweighted Unifrac)")
PcoA_plot_asvbu
#grid.arrange(PcoA_plot_asvbw, PcoA_plot_asvbu, ncol=2, nrow =1)

##6- PCoA Bray-Curtis - Bacteria ASV versus OTU
# PCoA ASV Bray-Curtis
spe <- read.csv2("asvb_tree.csv", row.names = 1)
env <- read.csv2("env.csv", row.names = 1)
spe.tr.asv <- decostand(spe, "hel")
dist.spe.asv <- vegdist(spe.tr.asv, method = "bray")
spe.pcoa <- cmdscale(dist.spe.asv,  k=(nrow(spe)-1), eig=TRUE)
pc1a <- round(spe.pcoa$eig[1]/sum(spe.pcoa$eig)*100)
pc2a <- round(spe.pcoa$eig[2]/sum(spe.pcoa$eig)*100)
pc3a <- round(spe.pcoa$eig[3]/sum(spe.pcoa$eig)*100)
pc4a <- round(spe.pcoa$eig[4]/sum(spe.pcoa$eig)*100)
spe.pcoa <- as.data.frame(spe.pcoa$points)
spe.pcoa <- spe.pcoa[,1:3]
group <- as.factor(env$Insect_Density_Class)
final <- cbind(spe.pcoa,group)
PcoA_plot_ASV <- ggplot(final, aes(x = V1, y = V2, colour = group, label =
row.names(spe))) + geom_point(size=8, show.legend = TRUE) +
#aes(shape=group), theme_light() + stat_ellipse() +
geom_hline(yintercept=0) + geom_vline(xintercept=0) + theme(legend.position
= "top")+ labs(x=paste("PCoA 1 (", pc1a, "%)", sep = ""), y=paste("PCoA 2
(", pc2a, "%)", sep = "", size=3))+ theme(panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA))+
ggtitle(label = "PCoA performed with bacteria, Bray-Curtis (ASV)")
PcoA_plot_ASV

# PCoA OTU Bray-Curtis
spe <- read.csv2("otu16b_tree.csv", row.names = 1)
env <- read.csv2("env.csv", row.names = 1)
spe.tr.otu <- decostand(spe, "hel")
dist.spe.otu <- vegdist(spe.tr.otu, method = "bray")
spe.pcoa <- cmdscale(dist.spe.otu,  k=(nrow(spe)-1), eig=TRUE)
pc1o <- round(spe.pcoa$eig[1]/sum(spe.pcoa$eig)*100)
pc2o <- round(spe.pcoa$eig[2]/sum(spe.pcoa$eig)*100)
pc3o <- round(spe.pcoa$eig[3]/sum(spe.pcoa$eig)*100)
pc4o <- round(spe.pcoa$eig[4]/sum(spe.pcoa$eig)*100)
spe.pcoa <- as.data.frame(spe.pcoa$points)
spe.pcoa <- spe.pcoa[,1:3]
group <- as.factor(env$Insect_Density_Class)
final <- cbind(spe.pcoa,group)
```

```
PcoA_plot_OTU <- ggplot(final, aes(x = V1, y = V2, colour = group, label =
row.names(spe))) + geom_point(size=8, show.legend = TRUE) +
#aes(shape=group), theme_light() + stat_ellipse() +
geom_hline(yintercept=0) + geom_vline(xintercept=0) + theme(legend.position
= "top") + labs(x=paste("PCoA 1 (", pc1o, "%)", sep = ""), y=paste("PCoA 2
(", pc2o, "%)", sep = "", size=3))+ theme(panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA)) +
ggtitle(label = "PCoA performed with bacteria, Bray-Curtis (OTU))")
PcoA_plot_OTU
#grid.arrange(PcoA_plot_ASV, PcoA_plot_OTU, ncol=2, nrow =1)

##7- Betadispersion - Permanova (Bray-Curtis) - Bacteria (ASV_OTU) - Fungi
# Bacteria ASV
spe <- read.csv2("asv16_trim1b.csv", row.names = 1)
env <- read.csv2("env.csv", row.names = 1)
spe.tr.asv <- decostand(spe, "hel")
dist.spe.asv <- vegdist(spe.tr.asv, method = "bray")
group <- as.factor(env$Insect_Density_Class)

# bétadispersion bacteria ASV
bdisper.asv <- betadisper(dist.spe.asv, group, "centroid")
bdisper.asv
anova(bdisper)

# Permanova bacteria ASV
PERMA_ASV <- adonis2(spe.tr.asv ~ group,  permutations = 9999, method =
"bray")
PERMA_ASV

# Bacteria OTU
spe <- read.csv2("otu16_trim1b.csv", row.names = 1)
env <- read.csv2("env.csv", row.names = 1)
spe.tr.otu <- decostand(spe, "hel")
dist.spe.otu <- vegdist(spe.tr.otu, method = "bray")
group <- as.factor(env$Insect_Density_Class)

# betadispersion bacteria OTU
bdisper.otu <- betadisper(dist.spe.otu, group, "centroid")
bdisper.otu
anova(bdisper)

# Permanova bacteria OTU
PERMA_OTU <- adonis2(spe.tr.otu ~ group,  permutations = 9999, method =
"bray")
PERMA_OTU

# Fungi ASV
spef <- read.csv2("asv16_trim1f.csv", row.names = 1)
env <- read.csv2("env.csv", row.names = 1)
group <- as.factor(env$Insect_Density_Class)
spe.trf <- decostand(spef, "hel") # for otu
dist.spef <- dist.ldc(spe.trf, method = "log.chord")

#betadispersion fungi ASV
bdisperf <- betadisper(dist.spef, group, "centroid")
bdisperf
anova(bdisperf)
# needs another transformation
spe.trf2 <- decostand(spef, "pa") # for otu
dist.spef2 <- dist.ldc(spe.trf2, method = "log.chord")
```

```
bdisperf2 <- betadisper(dist.spef2, group, "centroid")
bdisperf2
anova(bdisperf2)

# permanova fungi ASV
PERMA_Fungi <- adonis2(dist.spef2 ~ group, data = env, permutations = 9999,
method = "bray")
PERMA_Fungi

##8- Betadispersion - Permanova (Bray-Curtis) – Fungi
# bétadispersion OTU
spef <- read.csv2("asv16_trim1f.csv", row.names = 1)
env <- read.csv2("env.csv", row.names = 1)
group <- as.factor(env$Insect_Density_Class)
spe.trf <- decostand(spef, "hel") # for otu
dist.spef <- dist.ldc(spe.trf, method = "log.chord")
bdisperf <- betadisper(dist.spef, group, "centroid")
bdisperf
anova(bdisperf)
# needs another transformation
spe.trf2 <- decostand(spef, "pa") # for otu
dist.spef2 <- dist.ldc(spe.trf2, method = "log.chord")
bdisperf2 <- betadisper(dist.spef2, group, "centroid")
bdisperf2
anova(bdisperf2)
#boxplot(bdisperf2, shape=group, main = "Homogeneity of groups dispersion")

# permanova OTU
PERMA2 <- adonis2(dist.spef2 ~ group, data = env, permutations = 9999,
method = "bray")
#densityplot(permustats(PERMA2))
PERMA2

##9- Indicator species
spe <- read.csv2("asv16.csv", row.names = 1)
env <- read.csv2("env.csv", row.names = 1)
indval <- multipatt(as.data.frame(spe), func = "IndVal.g", group, duleg =
T, control = how(nperm=99999))
summary(indval)
```

5.3 Annexe 3 : Code R utilisé pour l'article # 2

```
##1- Import libraries
library(vegan)
library(ade4)
library(ape)
library(stats)
library(agricolae)
library(ggplot2)
library(gridExtra)
library(plyr)
library(plotly)
library(phyloseq)
library(microbiome)
library(clustsig)
library(ggdendro)
library(iNEXT)
library(adegenet)
library(phangorn)
library(adespatial)
library(indicspecies)
library(corrplot)
library(tidyverse)
library(dplyr)
library(MASS)
library(ellipse)
library(FactoMineR)
library(ggthemes)
library(SoDA)
library(adegraphics)
library(RColorBrewer)
library(scales)
library(vegetarian)
library(FD)
library(taxize)
library(picante)
library(ggpubr)
library(Polychrome)
library(rstatix)
library(gclus)
library(forecast)
library(VennDiagram)
library(missMDA)

#2- Trimming data
#Bacteria
otumat1 <- read.csv2("18asvb_raw.csv", row.names = 1)
otumatb <- otumat1[,-1]
taxmatb <- read.csv2("18taxab_raw.csv", row.names = 1)
otumatb <- as.matrix(otumatb)
taxmatb <- as.matrix(taxmatb)
class(otumatb)
class(taxmatb)
OTUb <- otu_table(t(otumatb), taxa_are_rows = TRUE)
TAXb <- tax_table(taxmatb)
dim(OTUb)
dim(TAXb)
phylob <- phyloseq(OTUb, TAXb)
require(genefilter)
flistb <- filterfun(pOverA(0.0005))
phylo1b <- filter_taxa(phylob, flistb, prune=TRUE)
```

```
phylo1b
bact2018 <- as.matrix(otu_table(phylo1b))
bact2018Taxa <- as.matrix(tax_table(phylo1b))
write.csv2(bact2018, file = "18asvb_trim.csv")
write.csv2(bact2018Taxa, file = "18taxab_trim.csv")
bact <- read.csv2("18asvb_trim.csv", row.names = 1)
dim(bact)

#Fungi
otumat1 <- read.csv2("18asvf_raw.csv", row.names = 1)
otumatf <- otumat1[,-1]
taxmatf <- read.csv2("18taxaf_raw.csv", row.names = 1)
otumatf <- as.matrix(otumatf)
taxmatf <- as.matrix(taxmatf)
class(otumatf)
class(taxmatf)
OTUf <- otu_table(t(otumatf), taxa_are_rows = TRUE)
TAXf <- tax_table(taxmatf)
dim(OTUf)
dim(TAXf)
phylof <- phyloseq(OTUf, TAXf) # phyloseq-class experiment-level object
require(genefilter)
flistf <- filterfun(pOverA(0.0005))
phylo1f <- filter_taxa(phylof, flistf, prune=TRUE)
phylo1f
fung2018 <- as.matrix(otu_table(phylo1f))
fung2018Taxa <- as.matrix(tax_table(phylo1f))
write.csv2(fung2018, file = "18asvf_trim.csv")
write.csv2(fung2018Taxa, file = "18taxaf_trim.csv")
fung <- read.csv2("18asvf_trim.csv", row.names = 1)
dim(fung)

##3- Microbiome gut and leaves
#Bacteria
otumatb1 <- read.csv2("18phylob.csv", row.names = 1)
taxmatb <- read.csv2("18taxab_trim_assigned2.csv", row.names = 1)
otumatb <- as.matrix(t(otumatb1))
taxmatb <- as.matrix(taxmatb)
OTU.bact <- otu_table(t(otumatb), taxa_are_rows = TRUE)
TAX.bact <- tax_table(taxmatb)
bact <- phyloseq(OTU.bact, TAX.bact)
bact.tr <- transform_sample_counts(bact, function(x) x / sum(x)*100 )
colB <-
c("#2966e7","#3e75e9","#5384eb","#6993ee","#7ea3f0","orange","#08bd5d","#20
c36d","#39ca7d","#52d08d","#ff4c4c","#ff6666","#8121f4","#ffff00","#7f858a"
,"#8B4B16")
Pb <- plot_bar(bact.tr, fill = "Class") + xlab("n=18/Biotope")+
ylab("Relative Abundance (%)")+ geom_bar(stat = "identity", position =
"stack") + scale_fill_manual(values = colB) + theme_classic()
phylaB <- Pb + theme(axis.text.x = element_text(face="bold", size=10),
axis.text.y = element_text(face="bold", size=10))+ theme(legend.title =
element_text(color="black", size=12, face="bold"),legend.position="right",
legend.text = element_text(colour="black", size=10, face="plain"))+
theme(axis.title.x = element_text(color="black", size=12, face="bold"),
axis.title.y = element_text(color="black", size=12, face="bold"))
phylaB
ggsave(filename = "bacteriomeF.tiff", height=5, width=5, units='in')
```

```
venn.bact <- draw.pairwise.venn(area1=392, area2=196, cross.area=186,
c("Guts", "Leaves"), scaled = T, fill=c("#f8766d", "#00ba38"), lty =
"blank", alpha = 0.30, cex = 2,     label.col = "black", cat.cex = 2,
cat.col = c("#f8766d", "#00ba38"),cat.pos  = c(285, 105), cat.dist = 0.09,
cat.just = list(c(-1, -1), c(1, 1)), ext.pos = 30, ext.dist = -0.05,
ext.length = 0.85, ext.line.lwd = 2, ext.line.lty = "dashed",col =
rep("#f8766d",2))
grid.draw(venn.bact)
tiff(filename = "VennBacteriaF.tiff", compression = "lzw+p", width = 400,
height = 500, units = "px"); grid.draw(venn.bact);

# Fungi
otumatf <- read.csv2("18phylof.csv", row.names = 1)
taxmatf <- read.csv2("18taxaf_trim_assigned2.csv", row.names = 1)
otumatf <- as.matrix(t(otumatf))
taxmatf <- as.matrix(taxmatf)
OTU.fung <- otu_table(t(otumatf), taxa_are_rows = TRUE)
TAX.fung <- tax_table(taxmatf)
fung <- phyloseq(OTU.fung, TAX.fung)
fung.tr <- transform_sample_counts(fung, function(x) x / sum(x)*100 )
colF <- c("#2966e7","#3e75e9","#5384eb","#6993ee","#7ea3f0",
"#94b2f3","#a9c1f5", "#bed1f7", "#d4e0fa","#e72966")
Pf <- plot_bar(fung.tr, fill = "Class") + xlab("n=18/Biotope")+
ylab("Relative Abundance (%)")+ geom_bar(stat = "identity", position =
"stack") + scale_fill_manual(values = colF) + theme_classic()
phylaf <- Pf + theme(axis.text.x = element_text(face="bold", size=10),
axis.text.y = element_text(face="bold", size=10))+ theme(legend.title =
element_text(color="black", size=12, face="bold"), legend.position="right",
legend.text = element_text(colour="black", size=10, face="plain")) +
theme(axis.title.x = element_text(color="black", size=12, face="bold"),
axis.title.y = element_text(color="black", size=12, face="bold"))
phylaf
ggsave(filename = "mycobiomeF.tiff", height=5, width=5, units='in')
venn.fung <- draw.pairwise.venn(area1=142, area2=173, cross.area=111,
c("Guts", "Leaves"), scaled = T, fill=c("#f8766d", "#00ba38"), lty =
"blank", alpha = 0.30, cex = 2,     label.col = "black", cat.cex = 2,
cat.col = c("#f8766d", "#00ba38"),cat.pos  = c(540, 505), cat.dist =
0.0001, cat.just = list(c(-0.3, 2), c(1, 1)), ext.pos = 2, ext.dist = 0.01,
ext.length = 0.05, ext.line.lwd = 2, ext.line.lty = "dashed",col =
rep("#f8766d",2))
grid.draw(venn.fung)
tiff(filename = "VennFungi.tiff", compression = "lzw+p", width = 400,
height = 500, units = "px"); grid.draw(venn.fung);
ggarrange(phylaf, NULL,venn.fung, nrow= 1,ncol = 3,labels = c("A","","B"),
heights = c(2,0.8,1.2), widths = c(2,0.8,1.2))
ggsave(filename = "Figure4bF.tiff", height=5, width=10, units='in')
ggarrange(phylaB,phylaf, nrow= 2,ncol = 1,align = "hv")
ggarrange(ggarrange(phylaB, NULL,venn.bact, nrow= 1,ncol = 3,labels =
c("A","","B"), heights = c(2,0.8,1.2), widths =
c(2,0.8,1.2)),NULL,ggarrange(phylaf, NULL,venn.fung, nrow= 1,ncol =
3,labels = c("C","","D"), heights = c(2,0.8,1.2), widths = c(2,0.8,1.2)),
nrow=3, heights = c(2,0.25,2))
ggsave(filename = "Figure4F.tiff", height=10, width=10, units='in')

#4- Permanova (Bray-Curtis) et dbRDA
# Bacteria
spe1 <- read.csv2("18asvb_trim.csv", row.names=1)
speb <- spe1[,-1]
group <- as.factor(spe1[,1])
speb.hel <- decostand(speb, "hel") # for otu
dist.speb <- vegdist(speb.hel, method = "bray")
```

```
speb.pcoa <- cmdscale(dist.speb, k=(nrow(speb)-1), eig=TRUE)
speb.pcoa <- as.data.frame(speb.pcoa$points)
dbrdaB <- dbrda(speb.pcoa~group,sqrt.dist = FALSE, add = FALSE,
dfun=vegdist, metaMDSdist = FALSE, na.action = na.fail, subset = NULL)
anova(dbrdaB)
RsquareAdj(dbrdaB)

# Fungi
spe1 <- read.csv2("18asvf_trim.csv", row.names=1)
spef <- spe1[,-1]
group <- as.factor(spe1[,1])
spef.hel <- decostand(spef, "hel") # for otu
dbrdaF <- capscale(spef.hel ~ spe1$group, distance = "bray")
anova(dbrdaF, permutations = how(nperm = 999), by="terms")
dbrdaF <- dbrda(spef.hel~group, distance = "bray", sqrt.dist = FALSE, add =
FALSE, dfun = vegdist, metaMDSdist = FALSE, na.action = na.fail, subset =
NULL)
anova(dbrdaF)
RsquareAdj(dbrdaF)

##5- Betadispersion - LCBD - SCBD - beta.div
# Bacteria
spe1 <- read.csv2("18asvb_trim.csv", row.names=1)
speb <- spe1[,-1]
group <- as.factor(spe1[,1])
speb.beta <- beta.div(speb, method = "hellinger", nperm = 9999)
summary(speb.beta)
speb.beta$beta # SSTotal and BDTotal
# Which species have a SCBD larger than the mean SCBD?
SCBD_bact <- speb.beta$SCBD
write.csv2(SCBD_bact, file = "SCBD_bactTotal.csv")
SCBD.largb <- sort(speb.beta$SCBD[speb.beta$SCBD >= mean(speb.beta$SCBD)])
SCBD.largb
ncol(SCBD.largb)
write.csv2(SCBD.largb, file = "SCBD_bact.csv")
speb.beta$LCBD # LCBD values
speb.beta$p.LCBD # p-values
p.adjust(speb.beta$p.LCBD, "holm") # Holm correction
row.names(speb[which(p.adjust(speb.beta$p.LCBD, "holm") <= 0.05),])

# Fungi
spe1 <- read.csv2("18asvf_trim.csv", row.names=1)
spef <- spe1[,-1]
group <- as.factor(spe1[,1])
spef.beta <- beta.div(spef, method = "hellinger", nperm = 9999)
summary(spef.beta)
spef.beta$beta # SSTotal and BDTotal
# Which species have a SCBD larger than the mean SCBD?
SCBD_fung <- spef.beta$SCBD
write.csv2(SCBD_fung, file = "SCBD_fungTotal.csv")
SCBD.largF <- spef.beta$SCBD[spef.beta$SCBD >= mean(spef.beta$SCBD)]
SCBD.largF
write.csv2(SCBD.largF, file = "SCBD_fung.csv")
spef.beta$LCBD # LCBD values
write.csv2(spef.beta$LCBD, file = "LCBD_fung.csv")
spef.beta$p.LCBD # p-values
p.adjust(spef.beta$p.LCBD, "holm") # Holm correction
row.names(spef[which(p.adjust(spef.beta$p.LCBD, "holm") <= 0.05),])
```

```
##6- Indicator species (gut and leaves)
spe1 <- read.csv2("18asvb_trim.csv", row.names = 1)
spe <- spe1[,-1]
group <- as.factor(spe1$group)
indvalB <- multipatt(as.data.frame(spe), func = "IndVal.g", group, duleg <-
T, control = how(nperm=9999))
summary(indvalB, indvalcom=T, alpha=0.001)
# A specificity/probable predictive value (exclusivity to the group)
# B fidelity/sensitivity (always faithful to the group)
spe1 <- read.csv2("18asvf_trim.csv", row.names = 1)
spe <- spe1[,-1]
group <- as.factor(spe1$group)
indvalF <- multipatt(as.data.frame(spe), func = "indval.g", group, duleg <-
T, control = how(nperm=9999))
summary(indvalF, indvalcom=T, alpha=0.001)


##7- Phytochemistry and microbial communities
# PCA to visualize the distribution of the samples
env1 <- read.csv2("18phytoF.csv", row.names = 1)
env <- env1[-11,] # remove NA
summary(env) # Descriptive statistics
env.stand <- decostand(env, "stand") # transform data
env.pca <- vegan::rda(env, scale = T) # compute pca
class(env.pca)
summary(env.pca) # Descriptive statistics, Default scaling 2
summary(env.pca, scaling = 1)
summary(env.pca, scaling = 1, axes=0) # remove some informations
source("cleanplot.pca.R")
par(mfrow = c(1, 2))
cleanplot.pca(env.pca, scaling = 1, mar.percent = 0.06)
cleanplot.pca(env.pca, scaling = 2, mar.percent = 0.06)

# Bacteria driven by phytochemistry ?
      # diversity indices
env1 <- read.csv2("18phytoF.csv", row.names = 1)
env <- env1[-11,] # remove NA
divb0 <- read.csv2("diversityB11.csv", row.names = 1)
divb <- divb0[1:18,-1]
divb <- divb[-11,]
divb.rda.all = rda(divb ~ ., data = env)
(adjR2 <- RsquareAdj (divb.rda.all)$adj.r.squared) # Not significant
#Try selection with ordistep()
mod1b <- rda(divb ~ 1, data = env)
step.forward1b <- ordistep(mod1b,scope = formula(divb.rda.all),
direction <- "forward", permutations = how(nperm = 999))
# No significant molecules

      # taxonomic structure
phypb1 <- read.csv2("18asvbG_phylum.csv", row.names = 1)
phycb1 <- read.csv2("18asvbG_class.csv", row.names = 1)
phypb1 <- phypb1[,-11]
phycb1 <- phycb1[,-22]
phypb <- cbind(phypb1,phycb1)
phypb <- phypb[-11,]
env <- read.csv2("18phytoF.csv", row.names = 1)
env <- env[-11,]
phypb.hel <- decostand(phypb, "hel")
phypb.rda.all <- rda(phypb.hel ~ ., data = env)
(adjR2 <- RsquareAdj (divb.rda.all)$adj.r.squared) # Not significant
```

```
mod2b <- rda(phypb.hel ~ 1, data = env)
step.forward2b <- ordistep(mod2b,scope = formula(phypb.rda.all),
direction <- "forward", permutations = how(nperm = 999))
# No significant molecules

      # SCBD
scbdb <- read.csv2("scbdB_asv.csv", row.names = 1)
env <- read.csv2("18phytoF.csv", row.names = 1)
env <- env[-11,]
scbdb <- scbdb[1:18,]
scbdb <- scbdb[-11,]
scbdb.hel <- decostand(scbdb, "hel")
scbdb.rda.all <- rda(scbdb.hel ~ .,data = env)
(R2adj <- RsquareAdj(scbdb.rda.all)$adj.r.squared) # interesting
#Check with ordistep()
mod3b <- rda(scbdb.hel ~ 1, data = env)
step.forward3b <- ordistep(mod3b,scope = formula(scbdb.rda.all),
direction <- "forward", permutations = how(nperm = 999))
# No significant molecules

##9- RDA & partitioning of the variation (gut bacterial community)
#      data preparation
source("triplot.rda.R")
spe <- read.csv2("18asvbG_raw30.csv", row.names=1)
spe1 <- read.csv2("18asvbL_raw30.csv", row.names=1)
spe2 <- read.csv2("18asvfL_raw30.csv", row.names=1)
env <- read.csv2("18phytoF.csv", row.names = 1)
geo <- read.csv2("GEO_18.csv", row.names = 1)
geo.xy <- geoXY(geo$lat, geo$lon)
geo.xy <- geo.xy[-11,]
spe <- spe[-11,]
spe1 <- spe1[-11,]
spe2 <- spe2[-11,]
env <- env[-11,]
# Hellinger-transform the species dataset
spe.hel <- decostand(spe, "hel")
spe1.hel <- decostand(spe1, "hel")
spe2.hel <- decostand(spe2, "hel")
env.stand <- decostand(env, "stand", na.rm = T)
geo.stand <- decostand(geo.xy,"stand")

      # Leaves' bacteria on gut's bacteria
spe.rda.all <- rda(spe.hel ~ ., data=spe1.hel, na.rm = T)
(R2a.all <- RsquareAdj(spe.rda.all)$adj.r.squared)
triplot.rda(spe.rda.all, site.sc="lc", scaling=1, cex.char2=0.8, pos.env=2,
mult.spe=1.1, mar.percent=-0.02)
mod1 <- rda(spe.hel ~ 1, data=spe1.hel, na.rm = T)
step1.forward <- ordistep(mod1,scope=formula(spe.rda.all),
direction="forward", permutations=how(nperm=9999))
RsquareAdj(step1.forward) # Some ASV have been selected
# 6 ASV selected with significant p-value
spe1.signif <- subset(spe1.hel, select =
c("Asvb105","Asvb869","Asvb78","Asvb299","Asvb205","Asvb5")) # 30
spe.rda.signif1 <- rda(spe.hel~., data=spe1.signif, na.rm = T)
summary(spe.rda.signif1, display=NULL)
(R2adj <- RsquareAdj(spe.rda.signif1)$adj.r.squared)
anova.cca(spe.rda.signif1, step=1000)
anova.cca(spe.rda.signif1, step=1000, by="axis")
anova.cca(spe.rda.signif1, permutations=how(nperm=999), by="terms")
vif.cca(spe.rda.signif1)
```

```
triplot.rda(spe.rda.signif1, site.sc="lc", scaling=1, plot.sites=FALSE,
label.sites=FALSE, cex.char2=0.8, pos.env=2, mult.spe=1.1, mar.percent=-
0.02)
ScoresBB1 <- scores(spe.rda.all, choices = 1:2, scaling = 1, display =
"species")
write.csv2(ScoresBB1, file = "ScoreBB_scal1.csv")
ScoresBB2 <- scores(spe.rda.all, choices = 1:2, scaling = 2, display =
"species")
write.csv2(ScoresBB2, file = "ScoreBB_scal2.csv")
```

```
      # Leaves' fungi on gut's bacteria
mod2 <- rda(spe.hel ~ 1, data=spe2)
spe.rda2.all <- rda(spe.hel ~ ., data=spe2.hel)
step2.forward <- ordistep(mod2, scope=formula(spe.rda2.all),
direction="forward", permutations=how(nperm=9999))
RsquareAdj(step2.forward)
# 6 ASV selected with significant p-value)
spe2.signif <- subset(spe2.hel, select =
c("Asvf26","Asvf145","Asvf235","Asvf90","Asvf263","Asvf173"))
spe.rda.signif2 <- rda(spe.hel~., data=spe2.signif)
summary(spe.rda.signif2, display=NULL)
(R2adj = RsquareAdj(spe.rda.signif2)$adj.r.squared)
anova.cca(spe.rda.signif2, step=1000)
anova.cca(spe.rda.signif2, step=1000, by="axis")
anova.cca(spe.rda.signif2, permutations=how(nperm=999), by="terms")
vif.cca(spe.rda.signif2) # VIF < 5
triplot.rda(spe.rda.signif2, site.sc="lc", scaling=1, plot.sites=FALSE,
label.sites=FALSE, cex.char2=0.8, pos.env=2, mult.spe=1.1, mar.percent=-
0.02)
ScoresFB1 <- scores(spe.rda.all, choices = 1:2, scaling = 1, display =
"species")
write.csv2(ScoresFB1, file = "ScoreFB_scal1.csv")
ScoresFB2 <- scores(spe.rda.all, choices = 1:2, scaling = 2, display =
"species")
write.csv2(ScoresFB2, file = "ScoreFB_scal2.csv")
```

```
      # metabolites on gut's bacteria
env.rda.all <- rda(spe.hel ~ ., data = env)
mod30 <- rda(spe.hel ~ 1, data=env)
step.forward <- ordistep(mod30,scope = formula(env.rda.all),
direction = "forward", permutations = how(nperm = 499))
RsquareAdj(step.forward)# No significants molecules
step3.forward <- ordistep(mod30, scope=formula(mod31), direction="both",
permutations=how(nperm=9999))
(adjR2 <- RsquareAdj (mod31)$adj.r.squared) # No metabolites selected
anova (mod31)  # no significant metabolites
```

```
      # geographical position on gut's bacteria
# plot a map draft with all the 17 sites
plot(geo.xy, asp=1)
text(geo.xy, labels=rownames(geo.xy), pos=3)
# Is there a significant linear spatial trend in the data?
trend.out <- rda(spe.hel, geo.xy)
anova(trend.out, step=1000, perm.max=1000)# Answer: no
# Compute model residual (detendenced data)
spe.lm <- lm(as.matrix(spe.hel) ~ as.matrix(geo.xy))
spe.resid <- residuals(spe.lm)
# Analyse dbMEM step by step
# First class of MEM model: dbMEM, threshold from minimum spanning tree
# Information used: links <= threshold, D values
# Construct dbMEM eigenfunctions
```

```
dxy <- dist(geo.xy)
spe.MEM <- dbmem(dxy) # this shows only two dbMEM ==> need to create a
spanning tree
# create a spanning tree
# need to compute the value of the threshold
#(min distance to keep all sites connected)
# Required library: adespatial
th <- give.thresh(dxy)
th
# Plot the minimum spaning tree
par(mfrow=c(1,1))
nb1 <- mst.nb(dxy)
nb1
wh1 <- which(as.matrix(dxy)==th,arr.ind=TRUE)
plot(nb1,geo.xy,pch=20,cex=2,lty=3)
#lines(geo.xy[wh1[1,],1],geo.xy[wh1[1,],2],lwd=2)
#title(main="Maximum distance of the minimum spanning tree in bold")
# Load the function plot.links() included in file "plot.links.R".
# Now,Plot the links on a map of the Gault Reserve
source("plot.links.R")
plot.links(geo.xy, thresh=th)
# MEM analysis, detrended data
spe.rda.resid <- rda(spe.resid, spe.MEM)
anova(spe.rda.resid, step=1000, perm.max=1000)# model not significant
spe.dbmem.tmp <- dbmem(geo.xy, silent = FALSE,MEM.autocor = "all",
thresh=th)
spe.dbmem <- as.data.frame(spe.dbmem.tmp)
## Step 1 Truncation distance used above:
(thr <- give.thresh(dist(geo.xy)))
# Display and count the eigenvalues
attributes(spe.dbmem.tmp)$values
length(attributes(spe.dbmem.tmp)$values)
## Step 2. Run the global dbMEM analysis on the detrended
## Hellinger-transformed spe data
spe.h.det <- resid(lm(as.matrix(spe.hel) ~ ., data =
as.data.frame(geo.xy)))
(spe.dbmem.rda <- rda(spe.h.det ~., spe.dbmem))
anova(spe.dbmem.rda) # not significant
## Step 3. if the R-square is significant, one should compute the adjusted
## R2 and run a forward selection of the dbmem variables
(spe.R2a <- RsquareAdj(spe.dbmem.rda)$adj.r.squared)
(spe.dbmem.fwd <- forward.sel(spe.h.det, as.matrix(spe.dbmem),
adjR2thresh = 1))
(nb.sig.dbmem <- nrow(spe.dbmem.fwd))
# Number of signif. dbMEM
# Identity of the significant dbMEM in increasing order
(dbmem.sign <- sort(spe.dbmem.fwd[ ,2]))
# Write the significant dbMEM to a new object
dbmem.red <- spe.dbmem[ ,c(dbmem.sign)]
write.csv2(dbmem.red, file = "dbMEM_signifB.csv")
## Step 4. New dbMEM analysis with significant dbMEM variables
## Adjusted R-square after forward selection: R2adj =
dbmem.red <- read.csv2("dbMEM_signifB.csv", row.names = 1)
(spe.dbmem.rda2 <- vegan::rda(spe.h.det ~ ., as.data.frame(dbmem.red)))
(spe.fwd.R2a <- RsquareAdj(spe.dbmem.rda2)$adj.r.squared)
anova(spe.dbmem.rda2)
(axes.test <- anova(spe.dbmem.rda2, by = "axis"))
# Number of significant axes
(nb.ax <- length(which(axes.test[ , ncol(axes.test)] <= 0.05)))
```

```
## Step 5. Plot the significant canonical axes
spe.rda2.axes <- scores(spe.dbmem.rda2, choices = c(1:nb.ax), display =
"lc", scaling = 1) par(mfrow = c(1,nb.ax))
for(i in 1:nb.ax){
sr.value(geo.xy, spe.rda2.axes[ ,i], sub = paste("RDA",i), csub = 2)
}
triplot.rda(spe.dbmem.rda2, site.sc="lc", scaling=1, plot.sites=FALSE,
label.sites=FALSE, cex.char2=0.8, pos.env=2, mult.spe=1.1, mar.percent=-
0.02)
```

 # Partitioning of variance with Varpart

```
part_bact3 <- varpart(spe.hel, spe1.signif,spe2.signif,dbmem.red)
plot(part_bact3, digits=3, Xnames = c("Bacteria (leaves)", "Fungi
(leaves)", "Geography (sites)"), bg=2:5, cutoff = -Inf, cex = 1)
#ggsave(filename = "VarPartBact.tiff", height=5, width=5, units='in')
part_bact3
anova.cca(rda(spe.hel,spe1.signif), step=1000)
anova.cca(rda(spe.hel,spe2.signif), step=1000)
anova.cca(rda(spe.hel,dbmem.red), step=1000)
anova.cca(rda(spe.hel,spe1.signif,spe2.signif), step=1000)
anova.cca(rda(spe.hel,spe1.signif,dbmem.red), step=1000)
anova.cca(rda(spe.hel,spe2.signif,dbmem.red), step=1000)
anova.cca(rda(spe.hel,spe1.signif,cbind(spe2.signif,dbmem.red)), step=1000)
anova.cca(rda(spe.hel,spe2.signif,cbind(spe1.signif,dbmem.red)), step=1000)
anova.cca(rda(spe.hel,dbmem.red,cbind(spe2.signif,spe1.signif)), step=1000)
anova.cca(rda(spe.hel,cbind(dbmem.red,spe1.signif,spe2.signif)), step=1000)
write.csv2(part_bact$part$fract, file = "varpartF_fract2.csv")
write.csv2(part_bact$part$indfract, file = "varpartF_indfract2.csv")
write.csv2(part_bact$part$contr1, file = "varpartF_contrl2.csv")
part_bact2 <- varpart(spe.hel, dbmem.red,spe2.signif)
plot(part_bact2, digits=3, Xnames = c("Bacteria (leaves)", "Fungi
(leaves)"), bg=3:5, cutoff = -Inf, cex = 1)
ggsave(filename = "VarPartBact.tiff", height=5, width=5, units='in')
part_bact2
```

##10- RDA & variance partitioning on gut fungal community

 # data preparation

```
source("triplot.rda.R")
spe <- read.csv2("18asvfG_raw30.csv", row.names=1)
spe1 <- read.csv2("18asvbL_raw30.csv", row.names=1)
spe2 <- read.csv2("18asvfL_raw30.csv", row.names=1)
env <- read.csv2("18phytoF.csv", row.names = 1)
geo <- read.csv2("Geo_18.csv", row.names = 1)
geo.xy <- geoXY(geo$lat, geo$lon)
geo.xy <- geo.xy[-11,]
spe <- spe[-11,]
spe1 <- spe1[-11,]
spe2 <- spe2[-11,]
env <- env[-11,]
# Hellinger-transform the species dataset
spe.hel <- decostand(spe, "hel")
spe1.hel <- decostand(spe1, "hel")
spe2.hel <- decostand(spe2, "hel")
env.stand <- decostand(env, "stand")
```

 # Leaves' bacteria on gut's fungi

```
spe.rda.all <- rda(spe.hel ~ ., data=spe1.hel)
(R2a.all <- RsquareAdj(spe.rda.all)$adj.r.squared)
triplot.rda(spe.rda.all, site.sc="lc", scaling=1, plot.sites=FALSE,
label.sites=FALSE, cex.char2=0.8, pos.env=2, mult.spe=1.1, mar.percent=-
0.02)
```

```
mod1 <- rda(spe.hel ~ 1, data=spe1.hel)
step1.forward <- ordistep(mod1,scope=formula(spe.rda.all),
direction="forward", permutations=how(nperm=999))
RsquareAdj(step1.forward)
# 7 ASV with significant p-value
spe1.signif <- subset(spe1.hel, select =
c("Asvb52","Asvb633","Asvb10","Asvb740","Asvb896","Asvb114","Asvb503"))
spe.rda.signif1 <- rda(spe.hel~., data=spe1.signif)
summary(spe.rda.signif1, display=NULL)
(R2adj <- RsquareAdj(spe.rda.signif1)$adj.r.squared)
anova.cca(spe.rda.signif1, step=1000)
anova.cca(spe.rda.signif1, step=1000, by="axis")
anova.cca(spe.rda.signif1, permutations=how(nperm=999), by="terms")
vif.cca(spe.rda.signif1)
triplot.rda(spe.rda.signif1, site.sc="lc", scaling=1, plot.sites=FALSE,
label.sites=FALSE, cex.char2=0.8, pos.env=2, mult.spe=1.1, mar.percent=-
0.02)
ScoresBF1 <- scores(spe.rda.all, choices = 1:2, scaling = 1, display =
"species")
write.csv2(ScoresBF1, file = "ScoreBF_scal1.csv")
ScoresBF2 <- scores(spe.rda.all, choices = 1:2, scaling = 2, display =
"species")
write.csv2(ScoresBF2, file = "ScoreBF_scal2.csv")

      # Leaves' fungi on gut's fungi
mod2 <- rda(spe.hel ~ 1, data=spe2.hel)
spe.rda2.all <- rda(spe.hel ~ ., data=spe2.hel)
step2.forward <- ordistep(mod2, scope=formula(spe.rda2.all),
direction="forward", #permutations=how(nperm=9999))
RsquareAdj(step2.forward)
#4 ASV with significant p-values
spe2.signif <- subset(spe2.hel, select =
c("Asvf486","Asvf292","Asvf8","Asvf123"))
spe.rda.signif2 <- rda(spe.hel~., data=spe2.signif)
summary(spe.rda.signif2, display=NULL)
(R2adj <- RsquareAdj(spe.rda.signif2)$adj.r.squared)
anova.cca(spe.rda.signif2, step=1000)
anova.cca(spe.rda.signif2, step=1000, by="axis")
anova.cca(spe.rda.signif2, permutations=how(nperm=999), by="terms")
vif.cca(spe.rda.signif2)
triplot.rda(spe.rda.signif2, site.sc="lc", scaling=1, plot.sites=FALSE,
label.sites=FALSE, cex.char2=0.8, pos.env=2, mult.spe=1.1, mar.percent=-
0.02)
ScoresFF1 <- scores(spe.rda.all, choices = 1:2, scaling = 1, display =
"species")
write.csv2(ScoresFF1, file = "ScoreFF_scal1.csv")
ScoresFF2 <- scores(spe.rda.all, choices = 1:2, scaling = 2, display =
"species")
write.csv2(ScoresFF2, file = "ScoreFF_scal2.csv")

      # Secondary metabolites on gut's fungi
mod3 <- rda(spe.hel ~ 1, data=env)
spe.rda3.all <- rda(spe.hel ~ ., data=env)
step3.forward <- ordistep(mod3, scope=formula(spe.rda3.all),
direction="forward", permutations=how(nperm=9999))
RsquareAdj(step3.forward)
spe3.signif <- subset(env.stand, select = c())
env.rda.signif = rda(spe.hel~., data=env.signif)
summary(env.rda.signif, display=NULL)
(R2adj = RsquareAdj(env.rda.signif)$adj.r.squared)
anova.cca(env.rda.signif, step=1000)


anova.cca(env.rda.signif, step=1000, by="axis")
anova.cca(env.rda.signif, permutations=how(nperm=999), by="terms")
```

```
      # geographical position on gut's bacteria
# plot a map draft with all the 17 sites
plot(geo.xy, asp=1)
text(geo.xy, labels=rownames(geo.xy), pos=3)
# Is there a significant linear spatial trend in the data
trend.out <- rda(spe.hel, geo.xy)
anova(trend.out, step=1000, perm.max=1000)# Answer: no
# Compute model residual (detendenced data)
spe.lm <- lm(as.matrix(spe.hel) ~ as.matrix(geo.xy))
spe.resid <- residuals(spe.lm)
# Analyse dbMEM par etapes
# create a spanning tree
# First class of MEM model: dbMEM, threshold from minimum spanning tree
# Information used: links <= threshold, D values
# Construct dbMEM eigenfunctions
# Required library:spacemakeR
dxy <- dist(geo.xy)
spe.MEM <- dbmem(dxy)
# Value of the threshold (min distance to keep all sites connected)?
# Required library: adespatial
th <- give.thresh(dxy)
th
# Plot the minimum spaning tree
par(mfrow=c(1,1))
nb1 <- mst.nb(dxy)
nb1
wh1 <- which(as.matrix(dxy)==th,arr.ind=TRUE)
plot(nb1,geo.xy,pch=20,cex=2,lty=3)
#lines(geo.xy[wh1[1,],1],geo.xy[wh1[1,],2],lwd=2)
#title(main="Maximum distance of the minimum spanning tree in bold")
# Load the function plot.links() that you will find in file "plot.links.R".
# Plot the links on a map of the Gault Reserve
source("plot.links.R")
plot.links(geo.xy, thresh=th)
# MEM analysis, detrended data
spe.rda.resid <- rda(spe.resid, spe.MEM)
anova(spe.rda.resid, step=1000, perm.max=1000)# model not significant
spe.dbmem.tmp <- dbmem(geo.xy, silent = FALSE,MEM.autocor = "all",
thresh=th)
spe.dbmem <- as.data.frame(spe.dbmem.tmp)
# Truncation distance used above:
(thr <- give.thresh(dist(geo.xy)))
# Step 1. Display and count the eigenvalues
attributes(spe.dbmem.tmp)$values
length(attributes(spe.dbmem.tmp)$values)
## Step 2. Run the global dbMEM analysis on the detrended
## Hellinger-transformed spe data
spe.h.det <- resid(lm(as.matrix(spe.hel) ~ ., data =
as.data.frame(geo.xy)))
(spe.dbmem.rda <- rda(spe.h.det ~., spe.dbmem))
anova(spe.dbmem.rda) # not significant
## Step 3. if the R-square is significant, one should compute the adjusted
## R2 and run a forward selection of the dbmem variables
(spe.R2a <- RsquareAdj(spe.dbmem.rda)$adj.r.squared)
(spe.dbmem.fwd <- forward.sel(spe.h.det, as.matrix(spe.dbmem),
adjR2thresh = 1))
(nb.sig.dbmem <- nrow(spe.dbmem.fwd))
write.csv2(spe.MEM, file = "dbMEMi_fung.csv")
# Number of signif. dbMEM
# Identity of the significant dbMEM in increasing order
(dbmem.sign <- sort(spe.dbmem.fwd[ ,2]))
```

```
# Write the significant dbMEM to a new object
dbmem.red <- spe.dbmem[ ,c(dbmem.sign)]
write.csv2(dbmem.red, file = "dbMEM_signifF.csv")
## Step 4. New dbMEM analysis with significant dbMEM variables
## Adjusted R-square after forward selection: R2adj =
dbmem.red <- read.csv2("dbMEM_signifF.csv", row.names = 1)
(spe.dbmem.rda2 <- vegan::rda(spe.h.det ~ ., as.data.frame(dbmem.red)))
(spe.fwd.R2a <- RsquareAdj(spe.dbmem.rda2)$adj.r.squared)
anova(spe.dbmem.rda2)
(axes.test <- anova(spe.dbmem.rda2, by = "axis"))
# Number of significant axes
(nb.ax <- length(which(axes.test[ , ncol(axes.test)] <= 0.05)))
## Step 5. Plot the significant canonical axes
spe.rda2.axes <- scores(spe.dbmem.rda2, choices = c(1:nb.ax), display =
"lc", scaling = 1) par(mfrow = c(1,nb.ax))
for(i in 1:nb.ax){sr.value(geo.xy, spe.rda2.axes[ ,i], sub =
paste("RDA",i), csub = 2)
}
triplot.rda(spe.dbmem.rda2, site.sc="lc", scaling=1, plot.sites=FALSE,
label.sites=FALSE, cex.char2=0.8, pos.env=2, mult.spe=1.1, mar.percent=-
0.02)
```

 # Partitioning of variance with Varpart

```
# For 3 explanatory tables
part_fung <- varpart(spe.hel, spe1.signif,spe2.signif,dbmem.red)
plot(part_fung, digits=3, Xnames = c("Bacteria (leaves)", "Fungi
(leaves)","Geography (sites)"), bg=2:4, cex = 1.5, cutoff = 0, id.size =
1.5)
ggsave(filename = "VarPartFung.tiff", height=5, width=5, units='in')
part_fung
anova.cca(rda(spe.hel,spe1.signif), step=1000)
anova.cca(rda(spe.hel,spe2.signif), step=1000)
anova.cca(rda(spe.hel,dbmem.red), step=1000)
anova.cca(rda(spe.hel,spe1.signif,spe2.signif), step=1000)
anova.cca(rda(spe.hel,spe1.signif,dbmem.red), step=1000)
anova.cca(rda(spe.hel,spe2.signif,dbmem.red), step=1000)
anova.cca(rda(spe.hel,spe1.signif,cbind(spe2.signif,dbmem.red)), step=1000)
anova.cca(rda(spe.hel,spe2.signif,cbind(spe1.signif,dbmem.red)), step=1000)
anova.cca(rda(spe.hel,dbmem.red,cbind(spe2.signif,spe1.signif)), step=1000)
anova.cca(rda(spe.hel,cbind(dbmem.red,spe1.signif,spe2.signif)), step=1000)
write.csv2(part_fung$part$fract, file = "varpartF_fract2.csv")
write.csv2(part_fung$part$indfract, file = "varpartF_indfract2.csv")
write.csv2(part_fung$part$contr1, file = "varpartF_contrl2.csv")
```

5.4 Annexe 4 : Code R utilisé pour l'article # 3

```
##1- Import librairies
library(vegan)
library(ggplot2)
library(gridExtra)
library(plyr)
library(dplyr)
library(microbiome)
library(plotly)
library(phyloseq)
library(clustsig)
library(ggdendro)
library(stats)
library(ade4)
library(ape)
library(iNEXT)
library(adegenet)
library(phangorn)
library(adespatial)
library(indicspecies)
library(corrplot)
library(tidyverse)
library(MASS)
library(ellipse)
library(FactoMineR)
library(ggthemes)
library(SoDA)
library(adegraphics)
library(vegan)
library(RColorBrewer)
library(scales)
library(vegetarian)
library(ggplot2)
library(FD)
library(taxize)
library(picante)
library(tidyverse)
library(ggpubr)
library(RColorBrewer)
library(Polychrome)
library(rstatix)
library(gclus)
library(plyr)
library(cooccur)
library(visNetwork)
library(shiny)
library(gridExtra)
library(grid)
library(lattice)
library(igraph)
library(devtools)
library(ggnetwork)
library(centiserve)
library(ggtree)
library(GGally)
library(ggnet)
library(sna)
library(intergraph)
```

```
##2- Phylogenetic analyses
# BACTERIA
commb = read.csv2("58asvB_trim21.csv", row.names = 1)
# standardization
commb <- decostand(commb, method = "hel")
# check total abundance in each sample
apply(commb, 1, sum)
#import the fasta file to built the tree
dnab <- read.dna("bact_muscle.fa", format="fasta")
Db <- dist.dna(dnab, model = "JC69")
length(Db)
treb <- nj(Db)
#is.rooted(treb)
#class(treb)
# root the tree before phylogenetic analysis
treb1 <- root(treb, outgroup = 3, resolve.root = TRUE)
is.rooted(treb1) # the tree need to be rooted
class(treb1)
identical(treb,treb1)
# check for mismatches/missing species
treb1 <- ladderize(treb1)
# plot our phylogeny (the cex argument makes the labels small enough to
# read)
#plot(treb1, cex = 0.6)
#title("NJ rooted Tree")
names(treb1) # check the tree file
treb1$tip.label[1:5]
# how many tips does our phylogeny have?
Ntip(treb1)
#write.nexus(treb1, file = "tree.bactG")
#treeB <- read.nexus("tree.bactG")
# the resulting object is a list with $phy and $comm elements.  replace our
# original data with the sorted/matched data
combinedb <- match.phylo.comm(treb1, commb)
metadata <- read.csv2("metadata.csv", row.names = 1)
metadata <- metadata[,1:3]
#We should also check whether our community data and metadata are in the
same order.
all.equal(rownames(commb), rownames(metadata))
# compare species richness between fescue and mixedgrass habitats
#boxplot(specnumber(commb) ~ metadata$biotope, ylab = "# of species")
# plot species accumulion curve across samples
#plot(specaccum(commb), xlab = "# of samples", ylab = "# of species")

# Phylogenetic diversity
# Calculate Faith's PD
#prunedTreeb <- prune.sample(commb,treb1) #to match community data
#plot(prunedTreeb, cex = 0.5)
#commb.pd <- pd(commb, treb1)
#head(commb.pd)
# export the data
#write.csv2(commb.pd, file = "PD_bact.csv")
# Plot Faith's PD by habitat
#boxplot(commb.pd$PD ~ metadata$biotope, xlab = "Biotope", ylab = "Faith's
PD")
# Test for PD differences among habitats
#shapiro.test(commb.pd$PD)# t.test can not be use due to the data
distribution
#wilcox.test(commb.pd$PD ~ metadata$biotope)
# Compare PD and species richness, the covariation should appear here
```

```
# Phylogenetic structure (MPD, MNTD, SESMPD and SESMNTD)
# convert phylogenety to a distance matrix
phy.distb <- cophenetic(treb1)
# calculate ses.mpd
# Standardized effect sizes describe the difference
# between average phylogenetic distances in the observed
# communities versus null communities generated with some
# randomization method, standardized by the standard deviation
# of phylogenetic distances in the null data:
# SESmetric <- ( Metricobserved - mean(Metricnull) ) / sd(Metricnull)
# mpd will calculate the mean pairwise distance between
# all species or individuals in each community
commb.sesmpd <- ses.mpd(commb, phy.distb, null.model = "phylogeny.pool",
abundance.weighted = FALSE, runs = 999)
#head(commb.sesmpd)
# calculate ses.mntd
# the mntd function calculates the mean nearest taxon distance,
# the average distance separating each species or individual
# in the community from its closest heterospecific relative.
commb.sesmntd <- ses.mntd(commb, phy.distb, null.model = "phylogeny.pool",
abundance.weighted = FALSE, runs = 999)
#head(commb.sesmntd)
# export the data
#phylo_bact <- cbind(commb.sesmpd, commb.sesmntd)
#write.csv2(phylo_bact, file = "phylo_bactF1.csv")
# Positive SES values (mpd.obs.z > 0) and high quantiles
# (mpd.obs.p > 0.95) indicate phylogenetic evenness, while
# negative SES values and low quantiles (mpd.obs.p < 0.05)
# indicate phylogenetic clustering, relative to the null model.
# MPD is generally thought to be more sensitive to tree-wide
# patterns of phylogenetic clustering and eveness, while
# MNTD is more sensitive to patterns of evenness and clustering
# closer to the tips of the phylogeny.
# The output includes the following columns:
# ntaxa - Number of taxa in community
# mpd.obs - Observed mpd in community
# mpd.rand.mean - Mean mpd in null communities
# mpd.rand.sd - Standard deviation of mpd in null communities
# mpd.obs.rank - Rank of observed mpd vs. null communities
# mpd.obs.z - Standardized effect size of mpd vs. null communities
(equivalent to -NRI)
# mpd.obs.p - P-value (quantile) of observed mpd vs. null communities (=
mpd.obs.rank / runs + 1)
# runs - Number of randomizations

# compare ses.mpd between habitats
#plot(commb.sesmpd$mpd.obs.z ~ metadata$biotope, xlab = "Biotope", ylab =
"SES(MPD)", main = "SES.MPD between biotopes")
#abline(h = 0, col = "gray")
#wilcox.test(commb.sesmpd$mpd.obs.z ~ metadata$biotope)
# compare ses.mntd between habitats
#plot(commb.sesmntd$mntd.obs.z ~ metadata$biotope, xlab = "Biotope", ylab =
"SES(MNTD)", main = "SES.MNTD between biotopes")
#abline(h = 0, col = "gray")
#wilcox.test(commb.sesmntd$mntd.obs.z ~ metadata$biotope)

# Calculate Faith's PD - FUNGI
commf = read.csv2("58asvF_trim21.csv", row.names = 1)
# Turn percent cover to relative abundance by dividing each value by sample
# total abundance
commf <- decostand(commf, method = "hel")
```

```
# check total abundance in each sample
apply(commf, 1, sum)
dnaf <- read.dna("fung_muscle.fa", format="fasta")
Df <- dist.dna(dnaf, model = "JC69")
length(Df)
tref <- nj(Df)
is.rooted(tref)
#class(tref)
tref1 <- root(tref, outgroup = 3, resolve.root = TRUE)
is.rooted(tref1)
#class(tref1)
#identical(tref,tref1)
#tref1 <- ladderize(tref1)
#plot(tref1, cex = 0.6)
#title("NJ rooted Tree")
names(tref1)
tref1$tip.label[1:5]
# how many tips does our phylogeny have?
Ntip(tref1)
#plot(tref1$edge)
#plot(tref1, cex = 0.5)
# check for mismatches/missing species
combinedf <- match.phylo.comm(tref1, commf)
# the resulting object is a list with $phy and $comm elements.  replace our
# original data with the sorted/matched data
metadata <- read.csv2("metadata.csv", row.names = 1)
metadata <- metadata[,1:3]
#We should also check whether our community data and metadata are in the
same order.
all.equal(rownames(commf), rownames(metadata))
#prunedTreef <- prune.sample(commf,tref1) #to match community data or trait
data
#plot(prunedTreef, cex = 0.5)
commf.pd <- pd(commf, tref1)

##3- phylogenetic structure
phy.distf <- cophenetic(tref1)
commf.sesmpd <- ses.mpd(commf, phy.distf, null.model = "phylogeny.pool",
abundance.weighted = FALSE, runs = 999)
head(commf.sesmpd)
# calculate ses.mntd
commf.sesmntd <- ses.mntd(commf, phy.distf, null.model = "phylogeny.pool",
abundance.weighted = FALSE,runs = 999)
head(commf.sesmntd)
#export data
phylo_fung <- cbind(commf.sesmpd, commf.sesmntd)
write.csv2(phylo_fung, file = "phylo_fungF1.csv")
# compare ses.mpd between habitats
plot(commf.sesmpd$mpd.obs.z ~ metadata$biotope, xlab = "Biotope", ylab =
"SES(MPD)", main = "SES.MPD between biotopes (insects and leaves)")
abline(h = 0, col = "gray")
wilcox.test(commf.sesmpd$mpd.obs.z ~ metadata$biotope)

# compare ses.mntd between habitats
mpd_mntdF <- read.csv2("mpd_mntdF2.csv",row.names = 1)
col <- c("#be99b2", "#99bea5")

      #PD bacteria
PDb <- ggboxplot(mpd_mntdF, x = "Biotope", y = "PD_bact", font.label =
list(size = 10),color = "grey", fill = col)
# Ajoutez des p-values sur les graphiques en box plot
```

```
stat.test <- mpd_mntdF %>% wilcox_test(PD_bact ~ Biotope) %>%
adjust_pvalue(method = "bonferroni") %>% add_significance("p.adj")
stat.test
stat.test <- stat.test %>% add_xy_position(x = "Biotope", dodge = 0.8)
PDb + stat_pvalue_manual(stat.test,  label = "p", tip.length = 0)
PDb1 <- PDb + stat_pvalue_manual(stat.test,  label = "{p.adj}
{p.adj.signif}", tip.length = 0) + scale_y_continuous(expand =
expansion(mult = c(0, 0.1)))+ labs(title=NULL,x="", y = "Faith PD
(bactéries)", cex=12)+ theme(axis.title.x = element_text(color="black",
size=10, face="bold"), axis.title.y = element_text(color="black", size=10,
face="bold"))+ scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
PDb1
```

#SR bacteria

```
SRb <- ggboxplot(mpd_mntdF, x = "Biotope", y = "SR_bact", font.label =
list(size = 10), color = "grey", fill = col)
# Ajoutez des p-values sur les graphiques en box plot
stat.test <- mpd_mntdF %>% wilcox_test(SR_bact ~ Biotope) %>%
adjust_pvalue(method = "bonferroni") %>% add_significance("p.adj")
stat.test
stat.test <- stat.test %>% add_xy_position(x = "Biotope", dodge = 0.8)
SRb + stat_pvalue_manual(stat.test,  label = "p", tip.length = 0)
SRb1 <- SRb + stat_pvalue_manual(stat.test,  label = "{p.adj}
{p.adj.signif}", tip.length = 0) + scale_y_continuous(expand =
expansion(mult = c(0, 0.1)))+ labs(title=NULL,x="", y = "Nombre d'espèces
(bactéries)", cex=12)+ theme(axis.title.x = element_text(color="black",
size=10, face="bold"), axis.title.y = element_text(color="black", size=10,
face="bold"))+ scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
SRb1
ggarrange(SRb1, PDb1)
```

#mpd bacteria

```
pmpdb <- ggboxplot(mpd_mntdF, x = "Biotope", y = "mpd_bact", font.label =
list(size = 8), color = "grey", fill = col)
# Ajoutez des p-values sur les graphiques en box plot
stat.test <- mpd_mntdF %>% wilcox_test(mpd_bact ~ Biotope) %>%
adjust_pvalue(method = "bonferroni") %>% add_significance("p.adj")
stat.test
stat.test <- stat.test %>% add_xy_position(x = "Biotope", dodge = 0.8)
pmpdb + stat_pvalue_manual(stat.test,  label = "p", tip.length = 0)
pmpdb1 <- pmpdb + stat_pvalue_manual(stat.test,  label = "{p.adj}
{p.adj.signif}", tip.length = 0) + scale_y_continuous(expand =
expansion(mult = c(0, 0.1)))+ labs(title=NULL,x="", y = "SES.MPD
(bactéries)", cex=12)+ theme(axis.title.x = element_text(color="black",
size=10, face="bold"), axis.title.y = element_text(color="black", size=10,
face="bold"))+ scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
pmpdb1
```

#mntd bacteria

```
pmntdb <- ggboxplot(mpd_mntdF, x = "Biotope", y = "mntd_bact", font.label =
list(size = 10), color = "grey", fill = col)
# Ajoutez des p-values sur les graphiques en box plot
stat.test <- mpd_mntdF %>% wilcox_test(mntd_bact ~ Biotope) %>%
adjust_pvalue(method = "bonferroni") %>% add_significance("p.adj")
stat.test
stat.test <- stat.test %>% add_xy_position(x = "Biotope", dodge = 0.8)
pmntdb + stat_pvalue_manual(stat.test,  label = "p", tip.length = 0)
```

```
pmntdb1 <- pmntdb + stat_pvalue_manual(stat.test,  label = "{p.adj}
{p.adj.signif}", tip.length = 0) + scale_y_continuous(expand =
expansion(mult = c(0, 0.1)))+ labs(title=NULL,x="", y = "SES.MNTD
(bactéries)", cex=12)+ theme(axis.title.x = element_text(color="black",
size=10, face="bold"), axis.title.y = element_text(color="black", size=10,
face="bold"))+ scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
pmntdb1
```

#PD fungi

```
PDf <- ggboxplot(pd, x = "Biotope", y = "PD_fung", color = "grey", fill =
col)
# Ajoutez des p-values sur les graphiques en box plot
stat.test <- pd %>% wilcox_test(PD_fung ~ Biotope) %>% adjust_pvalue(method
= "bonferroni") %>% add_significance("p.adj")
stat.test
stat.test <- stat.test %>% add_xy_position(x = "Biotope", dodge = 0.8)
PDf + stat_pvalue_manual(stat.test,  label = "p", tip.length = 0)
PDf1 <- PDf + stat_pvalue_manual(stat.test,  label = "{p.adj}
{p.adj.signif}", tip.length = 0) + scale_y_continuous(expand =
expansion(mult = c(0, 0.1)))+ labs(title=NULL,x="", y = "Faith PD
(champignons)", cex=12)+ theme(axis.title.x = element_text(color="black",
size=10, face="bold"), axis.title.y = element_text(color="black", size=10,
face="bold"))+ scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
PDf1
```

#SR fungi

```
SRf <- ggboxplot(pd, x = "Biotope", y = "SR_fung", color = "grey", fill =
col)
# Ajoutez des p-values sur les graphiques en box plot
stat.test <- pd %>% wilcox_test(SR_fung ~ Biotope) %>% adjust_pvalue(method
= "bonferroni") %>% add_significance("p.adj")
stat.test
stat.test <- stat.test %>% add_xy_position(x = "Biotope", dodge = 0.8)
SRf + stat_pvalue_manual(stat.test,  label = "p", tip.length = 0)
SRf1 <- SRf + stat_pvalue_manual(stat.test,  label = "{p.adj}
{p.adj.signif}", tip.length = 0) + scale_y_continuous(expand =
expansion(mult = c(0, 0.1)))+ labs(title=NULL,x="", y = "Nombre d'espèces
(champignons)", cex=12)+ theme(axis.title.x = element_text(color="black",
size=10, face="bold"), axis.title.y = element_text(color="black", size=10,
face="bold"))+ scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
SRf1
```

#mpd fungi

```
pmpdf <- ggboxplot(mpd_mntdF, x = "Biotope", y = "mpd_fung", font.label =
list(size = 10), color = "grey", fill = col)
# Ajoutez des p-values sur les graphiques en box plot
stat.test <- mpd_mntdF %>% wilcox_test(mpd_fung ~ Biotope) %>%
adjust_pvalue(method = "bonferroni") %>% add_significance("p.adj")
stat.test
stat.test <- stat.test %>% add_xy_position(x = "Biotope", dodge = 0.8)
pmpdf + stat_pvalue_manual(stat.test, label = "p", tip.length = 0)
pmpdf1 <- pmpdf + stat_pvalue_manual(stat.test,  label = "{p.adj}
{p.adj.signif}", tip.length = 0) + scale_y_continuous(expand =
expansion(mult = c(0, 0.1)))+ labs(title=NULL,x="", y = "SES.MPD
(champignons)", cex=12)+ theme(axis.title.x = element_text(color="black",
size=10, face="bold"), axis.title.y = element_text(color="black", size=10,
face="bold"))+ scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
pmpdf1
```

```
      #mntd fungi
pmntdf <- ggboxplot(mpd_mntdF, x = "Biotope", y = "mntd_fung", font.label =
list(size = 10), color = "grey", fill = col)
# Ajoutez des p-values sur les graphiques en box plot
stat.test <- mpd_mntdF %>% wilcox_test(mntd_fung ~ Biotope) %>%
adjust_pvalue(method = "bonferroni") %>% add_significance("p.adj")
stat.test
stat.test <- stat.test %>% add_xy_position(x = "Biotope", dodge = 0.8)
pmntdf + stat_pvalue_manual(stat.test,  label = "p", tip.length = 0)
pmntdf1 <- pmntdf + stat_pvalue_manual(stat.test,  label = "{p.adj}
{p.adj.signif}", tip.length = 0) + scale_y_continuous(expand =
expansion(mult = c(0, 0.1)))+ labs(title=NULL,x="", y = "SES.MNTD
(champignons)", cex=12)+ theme(axis.title.x = element_text(color="black",
size=10, face="bold"), axis.title.y = element_text(color="black", size=10,
face="bold"))+ scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
pmntdf1

      # combine all the plots in the same figure
ggarrange(SRb1, PDb1,pmpdb1, pmntdb1,SRf1, PDf1,pmpdf1, pmntdf1, ncol = 4,
nrow = 2,labels = c("A","","","","B","","",""), common.legend = TRUE)

      #save the figure
ggsave(filename = "phylogenic_structure.tiff", height=10, width=10,
units='in')

# NRI NTI bacteria
      #NRI bacteria
nrib <- ggboxplot(mpd_mntdF, x = "Biotope", y = "NRI_bact", font.label =
list(size = 10), color = "grey", fill = col)+geom_abline(slope = 0,colour =
"grey50", size = 1)
nrib1 <- nrib + scale_y_continuous(expand = expansion(mult = c(0, 0.1)))+
labs(title=NULL,x="", y = "NRI (bactéries)", cex=10)+ theme(axis.title.x =
element_text(color="black", size=10, face="bold"), axis.title.y =
element_text(color="black", size=10, face="bold"))+
scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
nrib1

      #NTI bacteria
ntib <- ggboxplot(mpd_mntdF, x = "Biotope", y = "NTI_bact", font.label =
list(size = 10), color = "grey", fill = col)+geom_abline(slope = 0,colour =
"grey50", size = 1)
ntib1 <- ntib + scale_y_continuous(expand = expansion(mult = c(0, 0.1)))+
labs(title=NULL,x="", y = "NTI (bactéries)", cex=10)+ theme(axis.title.x =
element_text(color="black", size=10, face="bold"), axis.title.y =
element_text(color="black", size=10, face="bold"))+
scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
ntib1

# NRI NTI fungi
      #NRI fungi
nrif <- ggboxplot(mpd_mntdF, x = "Biotope", y = "NRI_fung", font.label =
list(size = 10), color = "grey", fill = col)+geom_abline(slope = 0,colour =
"grey50", size = 1)
nrif1 <- nrif + scale_y_continuous(expand = expansion(mult = c(0, 0.1)))+
labs(title=NULL,x="", y = "NRI (champignons)", cex=10)+ theme(axis.title.x
= element_text(color="black", size=10, face="bold"), axis.title.y =
element_text(color="black", size=10, face="bold"))+
scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
nrif1
```

```
      #NTI fungi
ntif <- ggboxplot(mpd_mntdF, x = "Biotope", y = "NTI_fung", font.label =
list(size = 10), color = "grey", fill = col)+geom_abline(slope = 0,colour =
"grey50", size = 1)
ntif1 <- ntif + scale_y_continuous(expand = expansion(mult = c(0, 0.1)))+
labs(title=NULL,x="", y = "NTI (champignons)", cex=10)+ theme(axis.title.x
= element_text(color="black", size=10, face="bold"), axis.title.y =
element_text(color="black", size=10, face="bold"))+
scale_x_discrete(labels=c("Insectes", "Phyllosphère"))
ntif1
ggarrange(nrib1, ntib1,nrif1, ntif1, ncol = 2, nrow = 2,labels =
c("A","","B",""), common.legend = TRUE)
ggsave(filename = "figure_suppl1.tiff", height=7, width=5, units='in')

##4- Comparative beta diversity analyses: Bray-Curtis versus mntd distance
# Bacteria- Bray curtis
commb = read.csv2("58asvB_trim21.csv", row.names = 1)
metadata <- read.csv2("metadata.csv", row.names = 1)
metadata <- metadata[,1:3]
commb <- decostand(commb, method = "hel")
dist.speb <- vegdist(commb, method = "bray")
spe.commb <- cmdscale(sqrt(dist.speb),  k=(nrow(commb)-1), eig=TRUE)
pc1b1 <- round(spe.commb$eig[1]/sum(spe.commb$eig)*100)
pc2b1 <- round(spe.commb$eig[2]/sum(spe.commb$eig)*100)
#pc3 <- round(spe.commf$eig[3]/sum(spe.commf$eig)*100)
#pc4 <- round(spe.commf$eig[4]/sum(spe.commf$eig)*100)
spe.commb <- as.data.frame(spe.commb$points)
spe.commb <- spe.commb[,1:3]
#grp <- as.factor(spe1$group)
final1b <- cbind(spe.commb,metadata$biotope)
biotope <- as.factor(metadata$biotope)
pb1 <- ggplot(final1b, aes(x = V1, y = V2, colour = biotope, label = NULL))
+ geom_point(size=4, show.legend = TRUE)+ scale_colour_manual(values =
c("#be99b2", "#99bea5")) + #aes(shape=biotope), theme_light() +
theme(legend.title = element_text(color = "black", size = 10, face="bold"),
legend.text = element_text(color = "black"),legend.position="top")+
stat_ellipse() + geom_hline(yintercept=0) + geom_vline(xintercept=0) +
labs(title = "",x=paste("PCoA 1 (", pc1b1, "%)", sep = ""), y=paste("PCoA 2
(", pc2b1, "%)", sep = ""))+ theme(axis.title.x =
element_text(color="black", size=10, face="bold"), axis.title.y =
element_text(color="black", size=10, face="bold"),panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA))+
geom_text(aes(label = rownames(commb)), vjust = "inward", hjust = "inward",
colour="black", size=3)
#  ggtitle(label = "PCoA performed with bacterial data")
pb1
adonis(dist.speb ~ biotope, data = metadata)

# Bacteria -PCoA based on phylogenetic structure (functional trait beta
diversity)
commb <- decostand(commb, method = "hel")
# check total abundance in each sample
apply(commb, 1, sum)
dnab <- read.dna("bact_muscle.fa", format="fasta")
Db <- dist.dna(dnab, model = "JC69")
length(Db)
treb <- nj(Db)
treb1 <- root(treb, outgroup = 3, resolve.root = TRUE)
is.rooted(treb1)
treb1 <- ladderize(treb1)
#class(treb1)
```

```
names(treb1)
treb1$tip.label[1:5]
# how many tips does our phylogeny have?
Ntip(treb1)
# check for mismatches/missing species
combinedb <- match.phylo.comm(treb1, commb) # does not work
# the resulting object is a list with $phy and $comm elements. replace our
# original data with the sorted/matched data
#We should also check whether our community data and metadata are in the
same order.
all.equal(rownames(commb), rownames(metadata))
# convert phylogenety to a distance matrix
phy.distb <- cophenetic(treb1)
```

 # PCoA with mntd distance

```
commb.mntd.dist <- comdistnt(commb, phy.distb, abundance.weighted = TRUE)
spe.commb2 <- cmdscale(commb.mntd.dist,  k=(nrow(commb)-1), eig=TRUE)
pc1b2 <- round(spe.commb2$eig[1]/sum(spe.commb2$eig)*100)
pc2b2 <- round(spe.commb2$eig[2]/sum(spe.commb2$eig)*100)
#pc3 <- round(spe.commf$eig[3]/sum(spe.commf$eig)*100)
#pc4 <- round(spe.commf$eig[4]/sum(spe.commf$eig)*100)
spe.commb2 <- as.data.frame(spe.commb2$points)
spe.commb2 <- spe.commb2[,1:3]
#grp <- as.factor(spe1$group)
final2b <- cbind(spe.commb2,metadata$biotope)
biotope <- as.factor(metadata$biotope)
pb2 <- ggplot(final2b, aes(x = V1, y = V2, colour = biotope)) +
geom_point(size=4, show.legend = T)+ scale_colour_manual(values =
c("#be99b2", "#99bea5")) + #aes(shape=biotope), theme_light() +
theme(legend.title = element_text(color = "black", size = 10, face="bold"),
legend.text = element_text(color = "black"),legend.position="bottom")+
stat_ellipse() + geom_hline(yintercept=0) + geom_vline(xintercept=0) +
labs(title = "",x=paste("PCoA 1 (", pc1b2, "%)", sep = ""), y=paste("PCoA 2
(", pc2b2, "%)", sep = ""))+ theme(axis.title.x =
element_text(color="black", size=10, face="bold"), axis.title.y =
element_text(color="black", size=10, face="bold"),panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA))+
geom_text(aes(label = rownames(commb)), vjust = "inward", hjust = "inward",
colour="black", size=3)
#  ggtitle(label = "PCoA performed with bacterial data")
pb2
adonis(commb.mntd.dist ~ biotope, data = metadata)
```

 # PCoA with MPD distance

```
commb.mpd.dist <- comdist(commb, phy.distb, abundance.weighted = TRUE)
spe.commb3 <- cmdscale(commb.mpd.dist,  k=(nrow(commb)-1), eig=TRUE)
pc1b3 <- round(spe.commb3$eig[1]/sum(spe.commb3$eig)*100)
pc2b3 <- round(spe.commb3$eig[2]/sum(spe.commb3$eig)*100)
#pc3 <- round(spe.commb3$eig[3]/sum(spe.commb3$eig)*100)
#pc4 <- round(spe.commb3$eig[4]/sum(spe.commb3$eig)*100)
spe.commb3 <- as.data.frame(spe.commb3$points)
spe.commb3 <- spe.commb3[,1:3]
final3b <- cbind(spe.commb3,metadata$biotope)
biotope <- as.factor(metadata$biotope)
```

```
pb3 <- ggplot(final3b, aes(x = V1, y = V2, colour = biotope, label = NULL))
+geom_point(size=4, show.legend = F)+ scale_colour_manual(values =
c("#be99b2", "#99bea5")) + stat_ellipse(show.legend=F)+ theme_light() +
geom_hline(yintercept=0) + geom_vline(xintercept=0) + theme(legend.position
= "top")+ labs(x=paste("PCoA 1 (", pc1b3, "%)", sep = ""), y=paste("PCoA 2
(", pc2b3, "%)", sep = ""))+ theme(axis.title.x =
element_text(color="black", size=10, face="bold"), axis.title.y =
element_text(color="black", size=10, face="bold"),panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA))+
geom_text(aes(label = rownames(commb)), vjust = "inward", hjust = "inward",
colour="black", size=3)
pb3
adonis(commb.mpd.dist ~ biotope, data = metadata)
#ggarrange(pb1,pb2,pb3, labels = c("Bray-Curtis","MPD","MNTD"),ncol = 2,
nrow=2, hjust = -1,vjust = 1.5, common.legend = T)
#ggarrange(pb1,ggarrange(pb2, pb3, ncol = 1,nrow = 2, labels =
c("MPD","MNTD"), hjust = -0.5,vjust = 0.5),nrow = 1,ncol = 2, hjust = -
0.5,vjust = 0.2, common.legend = T, labels = "Bray-Curtis")
#ggsave(filename = "Figure3.2.tiff", height=10, width=15, units='in')
```

#Fungi – PcoA

```
commf = read.csv2("58asvF_trim21.csv", row.names = 1)
metadata <- read.csv2("metadata.csv", row.names = 1)
metadata <- metadata[,1:3]
commf <- decostand(commf, method = "hel")
dist.spef <- vegdist(commf, method = "bray")
spe.commf <- cmdscale(dist.spef, k=(nrow(commf)-1), eig=TRUE)
pc1f1 <- round(spe.commf$eig[1]/sum(spe.commf$eig)*100)
pc2f1 <- round(spe.commf$eig[2]/sum(spe.commf$eig)*100)
#pc3 <- round(spe.commf$eig[3]/sum(spe.commf$eig)*100)
#pc4 <- round(spe.commf$eig[4]/sum(spe.commf$eig)*100)
spe.commf <- as.data.frame(spe.commf$points)
spe.commf <- spe.commf[,1:3]
#grp = as.factor(spe1$group)
final1f <- cbind(spe.commf,metadata$biotope)
biotope <- as.factor(metadata$biotope)
pf1 <- ggplot(final1f, aes(x = V1, y = V2, colour = biotope, label = NULL))
+geom_point(size=4, show.legend = TRUE)+ scale_colour_manual(values =
c("#be99b2", "#99bea5")) + #aes(shape=biotope), theme_light() +
theme(legend.title = element_text(color = "black", size = 10, face="bold"),
legend.text = element_text(color = "black"),legend.position="top")+
stat_ellipse(show.legend=F) + geom_hline(yintercept=0) +
geom_vline(xintercept=0) + labs(x=paste("PCoA 1 (", pc1f1, "%)", sep = ""),
y=paste("PCoA 2 (", pc2f1, "%)", sep = ""))+ theme(axis.title.x =
element_text(color="black", size=10, face="bold"), axis.title.y =
element_text(color="black", size=10, face="bold"),panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA)) +
geom_text(aes(label = rownames(commb)), vjust = "inward", hjust = "inward",
colour="black", size=3)
#  ggtitle(label = "PCoA performed with bacterial data")
pf1
adonis(dist.spef ~ biotope, data = metadata)
ggarrange(pb1,pf1, nrow = 1,ncol = 2, common.legend = T, hjust = -1, labels
= c("Bacteria","Fungi"), align = "hv")
ggsave(filename = "Figure3.2.tiff", height=10, width=15, units='in')
```

#Fungi – PCoA based on Bray-Curtis, non phylogenetic beta diversity

```
# Turn percent cover to relative abundance by dividing each value by sample
# total abundance (Hellinger transformation)
commf <- decostand(commf, method = "hel")
# check total abundance in each sample
```

```
apply(commf, 1, sum)


dnaf <- read.dna("fung_muscle.fa", format="fasta")
Df <- dist.dna(dnaf, model = "JC69")
length(Df)
tref <- nj(Df)
tref1 <- root(tref, outgroup = 3, resolve.root = TRUE)
is.rooted(tref1)
class(tref1)
identical(tref,tref1)
tref1 <- ladderize(tref1)
#class(treb1)
names(tref1)
tref1$tip.label[1:5]
# how many tips does our phylogeny have?
Ntip(tref1)
# check for mismatches/missing species
combinedf <- match.phylo.comm(tref1, commf) # does not work
# the resulting object is a list with $phy and $comm elements.  replace our
# original data with the sorted/matched data
#We should also check whether our community data and metadata are in the
same order.
all.equal(rownames(commf), rownames(metadata))
# convert phylogenety to a distance matrix
phy.distf <- cophenetic(tref1)

      # Fungi - PCoA with phylogenetic MNTD beta diversity
# calculate functional trait MNTD beta diversity
#commb.mntd.traits.dist <- comdistnt(commb, trait.dist, abundance.weighted
= TRUE)
commf.mntd.dist <- comdistnt(commf, phy.distf, abundance.weighted = TRUE)
spe.commf2 <- cmdscale(commf.mntd.dist,  k=(nrow(commf)-1), eig=TRUE)
pc1f2 <- round(spe.commf2$eig[1]/sum(spe.commf2$eig)*100)
pc2f2 <- round(spe.commf2$eig[2]/sum(spe.commf2$eig)*100)
#pc3 <- round(spe.commf$eig[3]/sum(spe.commf$eig)*100)
#pc4 <- round(spe.commf$eig[4]/sum(spe.commf$eig)*100)
spe.commf2 <- as.data.frame(spe.commf2$points)
spe.commf2 <- spe.commf2[,1:3]
#grp <- as.factor(spe1$group)
final2f <- cbind(spe.commf2,metadata$biotope)
biotope <- as.factor(metadata$biotope)
pf2 <- ggplot(final2f, aes(x = V1, y = V2, colour = biotope, label = NULL))
+geom_point(size=4, show.legend = TRUE)+ scale_colour_manual(values =
c("#be99b2", "#99bea5")) + #aes(shape=group), theme_light() +
theme(legend.title = element_text(color = "black", size = 10, face="bold"),
legend.text = element_text(color = "black"),legend.position="bottom")+
geom_hline(yintercept=0) + geom_vline(xintercept=0) + labs(title =
"",x=paste("PCoA 1 (", pc2f2, "%)", sep = ""), y=paste("PCoA 2 (", pc2f2,
"%)", sep = ""))+ theme(axis.title.x = element_text(color="black", size=10,
face="bold"), axis.title.y = element_text(color="black", size=10,
face="bold"),panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA))+
geom_text(aes(label = rownames(commb)), vjust = "inward", hjust = "inward",
colour="black", size=3)
#  ggtitle(label = "PCoA performed with bacterial data")
pf2
#  ggtitle(label = "PCoA performed with bacterial data")
adonis(commf.mntd.dist ~ biotope, data = metadata)

      # Fungi - PCoA with phylogenetic MPD beta diversity
commf.mpd.dist <- comdist(commf, phy.distf, abundance.weighted = TRUE)
spe.commf3 <- cmdscale(commf.mpd.dist,  k=(nrow(commf)-1), eig=TRUE)
pc1f3 <- round(spe.commf3$eig[1]/sum(spe.commf3$eig)*100)
```

```
pc2f3 <- round(spe.commf3$eig[2]/sum(spe.commf3$eig)*100)
#pc3 <- round(spe.commf$eig[3]/sum(spe.commf$eig)*100)


#pc4 <- round(spe.commf$eig[4]/sum(spe.commf$eig)*100)
spe.commf3 = as.data.frame(spe.commf3$points)
spe.commf3 = spe.commf3[,1:3]
final3f <- cbind(spe.commf3,metadata$biotope)
biotope <- as.factor(metadata$biotope)
pf3 <- ggplot(final3f, aes(x = V1, y = V2, colour = biotope, label = NULL))
+geom_point(size=4, show.legend = TRUE)+ scale_colour_manual(values =
c("#be99b2", "#99bea5"))+ stat_ellipse() + theme_light() +
geom_hline(yintercept=0) + geom_vline(xintercept=0) + labs(x=paste("PCoA 1
(", pc1f3, "%)", sep = ""), y=paste("PCoA 2 (", pc2f3, "%)", sep = ""))+
theme(panel.background =
element_blank(),plot.background=element_rect(fill="white",colour=NA))+
geom_text(aes(label = rownames(commf)), vjust = "inward", hjust = "inward",
colour="black", size=3)
#  ggtitle(label = "PCoA performed with bacterial data")
pf3
adonis2(commf.mpd.dist ~ biotope, data = metadata)
ggarrange(pb2,pf2, nrow = 1,ncol = 2, common.legend = T, labels =
c("Bacteria","Fungi"), align = "hv")
ggsave(filename = "Figure3.3.tiff", height=10, width=15, units='in')

##5- cooccurrences networks
# Insects
# Find significant pairwise co-occurrences considering ASV occuring in at
least 30% of the gut samples
otumat.gut <- read.csv2("58asv_network30G_1F.csv", row.names = 1)
otumat.gutT <- decostand(otumat.gut, "pa")
otumat.gutG <- as.matrix(otumat.gutT)
cooG.gut2a <- cooccur(otumat.gutG, spp_names = T,thresh=T, prob = "comb")
summary(cooG.gut2a)
pair.attributes(cooG.gut2a)
pair.profile(cooG.gut2a)
p1.gut2 <- plot(cooG.gut2a)
p1.gut2  + ggtitle("ASVs Co-Occurrence Matrix (Insects)")
#ggsave(filename = "cooccurrences_gut1.tiff", height=8, width=12,
units='in')
p2.gut2 <- obs.v.exp(cooG.gut2a)
p2.gut2 + ggtitle("Observed-Expected ASVs co-occurrences (insects")+
theme_grey(base_size = 20)
#ggarrange(plot.gut1, plot.gut2, ncol = 2,nrow = 1)
#ggsave(filename = "cooccurrences_gut2.tiff", height=8, width=12,
units='in')
#probG2a <- print(prob.table(cooG.gut2a))
#write.csv2(probG2a, file = "cooccurG2F.csv")
#to extract information for one individual species,
#pair(mod = cooG.gut1, spp = "Asvf276")
#Printing this class returns only the significant interactions.
#We'll store the significant interactions to create our edge list for
visNetwork.
#print and save significant cooccurences
cooG.gut2b <- print(cooccur(otumat.gutG, spp_names = T,thresh=T, prob =
"comb"))
#write.csv2(cooG.gut2b, file = "cooccurG2_signifF.csv")
# Check sp1_name matches numeric label for species.
cooG.gut2b[, "sp1_name"] == rownames(otumat.gutG)[cooG.gut2b$sp1]
cooG.gut2b[, "sp2_name"] == rownames(otumat.gutG)[cooG.gut2b$sp2]
# Create a data frame of the nodes in the network with specific colors.
```

```
colorsg <-
c("#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#b
e99b2","#be99b2","#be99b2","#be99b2","#ebf53b","#be99b2","#be99b2","#be99b2
","#be99b2","#be99b2","#ebf53b","#be99b2","#be99b2","#be99b2","#be99b2","#b
e99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2
","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#b
e99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2
","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#b
e99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2","#be99b2
","#be99b2")
# make the node proportional to the degree computed ("value")
nodesg <- data.frame(id = 1:nrow(otumat.gutG),value =
c(6,33,11,12,24,25,33,8,25,27,7,37,17,29,18,24,5,38,32,24,27,29,32,14,29,21
,14,8,26,31,19,4,9,28,30,9,3,11,17,3,31,22,20,26,32,20,8,11,3,3,24,32,9,7,3
,3,7,7,2,5,4),label = rownames(otumat.gutG), color = colorsg, font.size =
25,shadow = T)
#write.csv2(nodesg, file = "nodesG2F.csv")
# Create an edges dataframe from the significant pairwise co-occurrences.
edgesg <- data.frame(from = cooG.gut2b$sp1, to = cooG.gut2b$sp2, color =
ifelse(cooG.gut2b$p_gt <= 0.05, "#95a1ab", "#d62d3e"),  dashes =
ifelse(cooG.gut2b$p_lt <= 0.05, TRUE, FALSE))
# Network visualization
networkgut <- visNetwork(nodesg, edgesg, width="100%",
height="600px",main="Cooccurrence network in the insects gut (n=29)")%>%
visIgraphLayout(layout = "layout_with_dh")
print(networkgut)
#class(networkgut)
# Save figure
visSave(networkgut, file = "networkG2F.html", background = "white")
#create a graph level centralization measure from the centrality scores of
the vertices.
netG <- graph_from_data_frame(d=edgesg, vertices=nodesg, directed=F)

      # Topology measures
edge_density(netG)
# 0.2945355 connectance
deg <- igraph::degree(netG)
deg <- sort(igraph::degree(netG))
clo <- sort(igraph::closeness(netG))
#measureg <- cbind(deg, clo)
#colnames(measureg) <- c("degree","closenness")
#write.csv2(measureg, file = "measureG.csv")
par(mfrow = c(1,2))
plot(igraph::degree(netG), typ = "h", ylab="Value (gut)", col = "tomato",
main = "Degree")
plot(igraph::closeness(netG), ylab="", typ = "h", col = "tomato", main =
"Closeness Centrality of vertices")
par(mfrow = c(1,1))
V(netG)$size <- deg*0.8 # modulate the node size
gut <- ggnet2(netG, label = rownames(otumat.gut), color = colorsg,
label.size = 3, edge.color = ifelse(cooG.gut2b$p_gt <= 0.05, "#95a1ab",
"#d62d3e"), edge.lty = ifelse(cooG.gut2b$p_lt <= 0.05, 6, 1), size =
"degree", size.cut = 3, mode = "kamadakawai")
gut

      # modularity measure
# greedy method (hiearchical, fast method)
c2 <- cluster_leading_eigen(netG)
#modularity(c2)
mod <- modularity(c2, directed=F)
class(mod)

c2$options$n
```

```
# plot communities with shaded regions
# modularity
deg <- igraph::degree(netG, mode="all")# mode="all" considers the intrants
and outputs degrees
V(netG)$size <- deg*0.8
lg <- layout_with_kk(netG)
lg <- norm_coords(lg, ymin=-1.5, ymax=1.5, xmin=-1.5, xmax=1.5)
plot(c2, netG,
vertex.label.cex=.8,vertex.label.color="black",vertex.label.font=1,
rescale=F, layout=l*0.6)
plot(c2,netG, rescale=F, layout=l*1)
c2$membership
write.csv2(c2$membership, file = "c2G.csv")
```

 # Phyllosphere

```
# Find significant pairwise co-occurrences considering indicator species
found in the leaf
otumat.leaf <- read.csv2("58asv_network30L_1F.csv", row.names = 1)
otumat.leafT <- decostand(otumat.leaf, "pa")
otumat.leaf <- as.matrix(otumat.leafT)
cooL.leaf2a <- cooccur(otumat.leaf, spp_names = T,thresh=T, prob = "comb")
summary(cooL.leaf2a)
p1.leaf2 <- plot(cooL.leaf2a)
p1.leaf2 + ggtitle("ASVs Co-Occurrence Matrix (phyllosphere)")
ggsave(filename = "cooccurrences_leaf1.tiff", height=8, width=12,
units='in')
p2.leaf2 <- obs.v.exp(cooL.leaf2a)
p2.leaf2 + ggtitle("Observed-Expected ASVs co-occurrences (phyllosphere)")+
theme_grey(base_size = 20)
ggsave(filename = "cooccurrences_leaf2.tiff", height=8, width=12,
units='in')
probL2a <- print(prob.table(cooL.leaf2a))
write.csv2(probL2a, file = "cooccurL2F.csv")
#Printing this class returns only the significant interactions. We'll store
the significant interactions to create our edge list for visNetwork.
cooL.leaf2b <- print(cooccur(otumat.leaf, spp_names = T,thresh=T, prob =
"comb"))
#print and save significant cooccurences
write.csv2(cooL.leaf2b, file = "cooccurL2_signifF.csv")
# Check sp1_name matches numeric label for species.
cooL.leaf2b[, "sp1_name"] == rownames(otumat.leaf)[cooL.leaf2b$sp1]
cooL.leaf2b[, "sp2_name"] == rownames(otumat.leaf)[cooL.leaf2b$sp2]
# Create a data frame of the nodes in the network.
colorsl <-
c("#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#9
9bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5
","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#9
9bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5
","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#9
9bea5","#99bea5","#99bea5","#ebf53b","#99bea5","#99bea5","#99bea5","#99bea5
","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#ebf53b","#e
bf53b","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5
","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#9
9bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5
","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#9
9bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5
","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#9
9bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5
","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#9
9bea5","#99bea5","#99bea5","#99bea5","#99bea5","#ebf53b","#99bea5","#99bea5
","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#99bea5","#9
9bea5","#99bea5")
```

```
nodesl <- data.frame(id = 1:nrow(otumat.leaf), value =
c(4,14,7,14,3,11,11,6,12,7,12,4,14,3,3,13,3,1,14,17,16,13,15,8,15,9,11,15,1
8,6,15,16,21,7,17,11,6,13,5,21,26,6,9,10,7,14,14,12,11,9,14,10,16,4,3,15,2,
7,3,9,9,7,10,7,3,4,3,15,1,6,13,13,8,8,3,14,8,11,7,5,12,9,12,1,11,5,14,4,13,
6,10,5,6,4,4,8,4,11,4,7,10,5,6,7,15,5,7,10,6,5,4,7,3,10,13,8,4,16,10,4,15,4
,8,12,6,8,10,3,12),label = rownames(otumat.leaf), color = colorsl,
font.size = 25,shadow = T)
#write.csv2(nodesl, file = "nodesL2F.csv")
# Create an edges dataframe from the significant pairwise co-occurrences.
edgesl <- data.frame(from = cooL.leaf2b$sp1, to = cooL.leaf2b$sp2, color =
ifelse(cooL.leaf2b$p_gt <= 0.05, "#95a1ab", "#d62d3e"),  dashes =
ifelse(cooL.leaf2b$p_lt <= 0.05, TRUE, FALSE))
# Network visualization
networkleaf <- visNetwork(nodesl, edgesl, width="100%",
height="600px",main="Cooccurrences network in phyllosphere (n=29)")%>%
visNodes(shadow = list(enabled = F, size = 30))%>% visIgraphLayout(layout =
"layout_with_dh")
print(networkleaf)
print(networkgut)
#create a graph level centralization measure from the centrality scores of
the vertices.
netL <- graph_from_data_frame(d=edgesl, vertices=nodesl, directed=F)

      # Topology measures
edge_density(netL)
# 0.2945355 connectance
degL <- igraph::degree(netL)
degL <- sort(igraph::degree(netL))
cloL <- sort(igraph::closeness(netL))
#measureg <- cbind(deg,clo)
#colnames(measureg) <- c("degree","closenness")
#write.csv2(measureg, file = "measureG.csv")
par(mfrow = c(1,2))
plot(igraph::degree(netL), typ = "h", ylab="Value (phyllosphere)", col =
"tomato", main = "Degree")
plot(igraph::closeness(netL), ylab="", typ = "h", col = "tomato", main =
"Closeness Centrality of vertices")
par(mfrow = c(1,1))
V(netL)$size <- degL*0.8 # modulate the node size
leaf <- ggnet2(netL, label = rownames(otumat.leaf), color = colorsl,
label.size = 3, edge.color = ifelse(cooL.leaf2b$p_gt <= 0.05, "#95a1ab",
"#d62d3e"), edge.lty = ifelse(cooL.leaf2b$p_lt <= 0.05, 6, 1), size =
"degree", size.cut = 3, mode = "kamadakawai")
leaf

      # modularity measure
c2L = cluster_leading_eigen(netL)
#modularity(c2L)
modularity(c2L, directed=F)
# plot communities with shaded regions
degL <- degree(netL)
V(netL)$size <- degL*0.5
lf <- layout_with_kk(netL)
lf <- norm_coords(lf, ymin=-1.5, ymax=1.5, xmin=-1.5, xmax=1.5)
plot(c2L, netL,
vertex.label.cex=.8,vertex.label.color="black",vertex.label.font=1,
rescale=F, layout=l*0.6)
#plot(c2L,netL, rescale=F, layout=l*1)
#c2L$membership
#write.csv2(c2L$membership, file = "c2L.csv")
```

```
      # plots with ggnet
V(netL)$size <- degL*0.8
l <- layout_randomly(netL)
ggnet2(netL, label = rownames(otumat.leaf), color = colorsl, label.size =
3, edge.color = ifelse(cooL.leaf2b$p_gt <= 0.05, "#95a1ab", "#d62d3e"),
edge.lty = ifelse(cooL.leaf2b$p_lt <= 0.05, 6, 1), size = "degree",
size.cut = 3)
ggsave(filename = "Network1_phyllosphere.tiff", height=8, width=8,
units='in' )
ggnet2(netG, label = rownames(otumat.gut), color = colorsg, label.size = 3,
edge.color = ifelse(cooG.gut2b$p_gt <= 0.05, "#95a1ab", "#d62d3e"),
edge.lty = ifelse(cooG.gut2b$p_lt <= 0.05, 6, 1), size = "degree", size.cut
= 3)
ggsave(filename = "Network1_gut.tiff", height=8, width=8, units='in' )
```

BIBLIOGRAPHIE COMPILÉE

Adams AS, Aylward FO, Adams SM, Erbilgin N, Aukema BH, Currie CR, Suen G, Raffa KF (2013) Mountain pine beetles colonizing historical and naive host trees are associated with a bacterial community highly enriched in genes contributing to terpene metabolism. *Applied and environmental microbiology* 79(11):3468-3475.

AFSQ (2018) La clé forestière: Le frêne blanc. https://afsq.org/information-foret/nos-arbres/frene-blanc/.

Agler MT, Ruhe J, Kroll S, Morhenn C, Kim S-T, Weigel D, Kemen EM (2016) Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS biology* 14(1):e1002352.

Anulewicz AC (2006) Field surveys of emerald ash borer (Agrilus planipennis Fairmaire)(Coleoptera: Buprestidae) density and canopy dieback in three North American ash species. *The 2006 ESA Annual Meeting, December 10-13, 2006.*

Arimura G-i, Matsui K, Takabayashi J (2009) Chemical and molecular ecology of herbivore-induced plant volatiles: proximate factors and their ultimate functions. *Plant and Cell Physiology* 50(5):911-923.

Aukema JE, Leung B, Kovacs K, Chivers C, Britton KO, Englin J, Frankel SJ, Haight RG, Holmes TP, Liebhold AM (2011) Economic impacts of non-native forest insects in the continental United States. *PLoS One* 6(9):e24587.

Banerjee S, Kirkby CA, Schmutter D, Bissett A, Kirkegaard JA, Richardson AE (2016) Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. *Soil Biology and Biochemistry* 97:188-198.

Baranchikov Y, Mozolevskaya E, Yurchenko G, Kenis M (2008) Occurrence of the emerald ash borer, Agrilus planipennis in Russia and its potential impact on European forestry. *EPPO bulletin* 38(2):233-238.

Beauchemin NJ, Furnholm T, Lavenus J, Svistoonoff S, Doumas P, Bogusz D, Laplaze L, Tisa LS (2012) Casuarina root exudates alter the physiology, surface properties, and plant infectivity of Frankia sp. strain CcI3. *Appl. Environ. Microbiol.* 78(2):575-580.

Bellamy CL (2002) *Coleoptera: Buprestoidea.* CSIRO PUBLISHING,

Ben Guerrero E, Soria M, Salvador R, Ceja-Navarro JA, Campos E, Brodie EL, Talia P (2016) Effect of Different Lignocellulosic Diets on Bacterial Microbiota and Hydrolytic Enzyme Activities in the Gut of the Cotton Boll Weevil (Anthonomus grandis). *Frontiers in Microbiology* 7:2093.

Bentz BJ & Six DL (2006) Ergosterol content of fungi associated with Dendroctonus ponderosae and Dendroctonus rufipennis (Coleoptera: Curculionidae, Scolytinae). *Annals of the Entomological Society of America* 99(2):189-194.

Berasategui A, Axelsson K, Nordlander G, Schmidt A, Borg-Karlson AK, Gershenzon J, Terenius O, Kaltenpoth M (2016) The gut microbiota of the pine weevil is similar across Europe and resembles that of other conifer-feeding beetles. *Molecular ecology* 25(16):4014-4031.

Berg G, Rybakova D, Fischer D, Cernava T, Vergès M-CC, Charles T, Chen X, Cocolin L, Eversole K, Corral GH (2020) Microbiome definition re-visited: old concepts and new challenges. *Microbiome* 8(1):1-22.

Bergeron A (2016) *Caractérisation de la communauté bactérienne associée à l'agrile du frêne, Agrilus planipennis Fairmaire (Coleoptera: Buprestidae)*. (Université du Québec, Institut National de la Recherche Scientifique).

Bi JL, Toscano NC, Madore MA (2003) Effect of Urea Fertilizer Application on Soluble Protein and Free Amino Acid Content of Cotton Petioles in Relation to Silverleaf Whitefly (Bemisia argentifolii) Populations. *Journal of Chemical Ecology* 29(3):747-761.

Biedermann PH & Vega FE (2020) Ecology and evolution of insect–fungus mutualisms. *Annual Review of Entomology* 65:431-455.

Blanchet FG, Legendre P, Borcard D (2008) Forward selection of explanatory variables. *Ecology* 89(9):2623-2632.

Bodilis J, Nsigue-Meilo S, Besaury L, Quillet L (2012) Variable copy number, intra-genomic heterogeneities and lateral transfers of the 16S rRNA gene in Pseudomonas. *PloS one* 7(4):e35647.

Bohan D, Dumbrell A, Massol F (2017) *Networks of Invasion: Empirical Evidence and Case Studies.* Academic Press,

Bonaventure G, VanDoorn A, Baldwin IT (2011) Herbivore-associated elicitors: FAC signaling and metabolism. *Trends in plant science* 16(6):294-299.

Boone CK, Keefover-Ring K, Mapes AC, Adams AS, Bohlmann J, Raffa KF (2013) Bacteria associated with a tree-killing insect reduce concentrations of plant defense compounds. *Journal of Chemical Ecology* 39(7):1003-1006.

Borcard D, Gillet F, Legendre P (2018) *Numerical ecology with R.* Springer,

Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology* 73(3):1045-1055.

Bozorov TA, Rasulov BA, Zhang D (2019) Characterization of the gut microbiota of invasive Agrilus mali Matsumara (Coleoptera: Buprestidae) using high-throughput sequencing: uncovering plant cell-wall degrading bacteria. *Scientific Reports* 9(1):4923.

Briatte F (2021) ggnet: Functions to plot networks with ggplot2.

Broderick G (2004) Effect of low level monensin supplementation on the production of dairy cows fed alfalfa silage. *Journal of Dairy Science* 87(2):359-368.

Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR (2016) Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. *PLOS Biology* 14(11):e2000225.

Brownlie JC & Johnson KN (2009) Symbiont-mediated protection in insect hosts. *Trends in microbiology* 17(8):348-354.

Burr SJ & McCullough DG (2014) Condition of green ash (Fraxinus pennsylvanica) overstory and regeneration at three stages of the emerald ash borer invasion wave. *Canadian Journal of Forest Research* 44(7):768-776.

Butin H (1992) Effect of endophytic fungi from oak (Quercus robur L.) on mortality of leaf inhabiting gall insects. *European Journal of Forest Pathology* 22(4):237-246.

Cadotte MW, Jonathan Davies T, Regetz J, Kembel SW, Cleland E, Oakley TH (2010) Phylogenetic diversity metrics for ecological communities: integrating species richness, abundance and evolutionary history. *Ecology letters* 13(1):96-105.

Cale JA, Taft S, Najar A, Klutsch JG, Hughes CC, Sweeney JD, Erbilgin N (2016) Mountain pine beetle (Dendroctonus ponderosae) can produce its aggregation pheromone and complete brood development in naïve red pine (Pinus resinosa) under laboratory conditions. *Canadian Journal of Forest Research* 45(12):1873-1877.

Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal* 11(12):2639.

Cambon M, Ogier J-C, Lanois A, Ferdy J-B, Gaudriault S (2018) Changes in rearing conditions rapidly modify gut microbiota structure in Tenebrio molitor larvae. *bioRxiv* 10.1101/423178:423178.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 7(5):335-336.

Cappaert D, McCullough DG, Poland TM, Siegert NW (2005a) Emerald ash borer in North America: a research and regulatory challenge. *American Entomologist. 51 (3): 152-165.* 51(3).

Cappaert D, McCullough DG, Poland TM, Siegert NW (2005b) Emerald ash borer in North America: a research and regulatory challenge.

Cardoza YJ, Klepzig KD, Raffa KF (2006) Bacteria in oral secretions of an endophytic insect inhibit antagonistic fungi. *Ecological Entomology* 31(6):636-645.

Carson W & Schnitzer S (2011) *Tropical forest community ecology.* John Wiley & Sons,

CFIA (2020) Emerald Ash Borer - Agrilus planipennis. *Canadian Food Inspection Agency Canada*.

Chakraborty S, Whitehill JGA, Hill AL, Opiyo SO, Cipollini DON, Herms DA, Bonello P (2014) Effects of water availability on emerald ash borer larval performance and phloem phenolics of Manchurian and black ash. *Plant, Cell & Environment* 37(4):1009-1021.

Chararas C & Chipoulet J-M (1983) Studies on the digestion of cellulose by the larvae of the Eucalyptus borer, Phoracantha semipunctata (Coleoptera: Cerambycidae). *Australian journal of biological sciences* 36(3):223-234.

Chatterjee S, Datta S, Sharma S, Tiwari S, Gupta DK (2017) Health and environmental applications of gut microbiome: a review. *Ecological Chemistry and Engineering S* 24(3):467-482.

Chattopadhyay P, Banerjee G, Mukherjee S (2017) Recent trends of modern bacterial insecticides for pest control practice in integrated crop management system. *3 Biotech* 7(1):60.

Chen Y & Poland TM (2009) Interactive influence of leaf age, light intensity, and girdling on green ash foliar chemistry and emerald ash borer development. *J Chem Ecol* 35(7):806-815.

Chen Y, Whitehill JG, Bonello P, Poland TM (2011) Differential response in foliar chemistry of three ash species to emerald ash borer adult feeding. *Journal of chemical ecology* 37(1):29-39.

Clauset A, Newman ME, Moore C (2004) Finding community structure in very large networks. *Physical review E* 70(6):066111.

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42(Database issue):D633-642.

Colman DR, Toolson EC, Takacs-Vesbach C (2012) Do diet and taxonomy influence insect gut bacterial communities? *Molecular ecology* 21(20):5124-5137.

Comeau AM, Li WKW, Tremblay J-É, Carmack EC, Lovejoy C (2011) Arctic Ocean Microbial Community Structure before and after the 2007 Record Sea Ice Minimum. *PLOS ONE* 6(11):e27492.

Csardi G & Nepusz T (2006) The igraph software package for complex network research. *InterJournal, complex systems* 1695(5):1-9.

Davis TS, Hofstetter RW, Foster JT, Foote NE, Keim P (2011) Interactions between the yeast Ogataea pini and filamentous fungi associated with the western pine beetle. *Microbial Ecology* 61(3):626-634.

De Cáceres M & Jansen F (2016) Package 'indicspecies': relationship between species and groups of sites. R Package Version 1.7. 6.).

De Cáceres M & Jansen F (2020) Package 'indicspecies': relationship between species and groups of sites. R Package Version 1.7. 9.

De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* 119(10):1674-1684.

Delalibera I, Vasanthakumar A, Klepzig KD, Raffa KF (2007) Composition of the bacterial community in the gut of the pine engraver, Ips pini (Say)(Coleoptera) colonizing red pine. Symbiosis. *Symbiosis* 43:97-104.

Dillon R & Dillon V (2004) The gut bacteria of insects: nonpathogenic interactions. *Annual Reviews in Entomology* 49(1):71-92.

Douglas AE (2009) The microbial dimension in insect nutritional ecology. *Functional Ecology* 23(1):38-47.

Douglas AE (2013) Microbial brokers of insect-plant interactions revisited. *Journal of chemical ecology* 39(7):952-961.

Douglas AE (2015) Multiorganismal insects: diversity and function of resident microorganisms. *Annual review of entomology* 60:17-34.

Dray S, Bauman D, Blanchet G, Borcard D, Clappe S, Guenard G (2020) & Wagner, HH (2020). adespatial: Multivariate multiscale spatial analysis (R package version 0.3-8).).

Duan JJ, Bauer LS, Hansen JA, Abell KJ, Van Driesche R (2012) An improved method for monitoring parasitism and establishment of Oobius agrili (Hymenoptera: Encyrtidae), an egg parasitoid introduced for biological control of the emerald ash borer (Coleoptera: Buprestidae) in North America. *Biological Control* 60(3):255-261.

Duan JJ, Bauer LS, Van Driesche RG (2017) Emerald ash borer biocontrol in ash saplings: The potential for early stage recovery of North American ash trees. *Forest Ecology and Management* 394:64-72.

Durand A-A, Bergeron A, Constant P, Buffet J-P, Déziel E, Guertin C (2015) Surveying the endomicrobiome and ectomicrobiome of bark beetles: The case of Dendroctonus simplex. 5:17190.

Edgar R (2010a) *Usearch.* Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), p

Edgar RC (2010b) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19):2460-2461.

Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv* :081257.

Edgar RC & Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 31(21):3476-3482.

Emerald ash Borer information network (2020) *Cooperative Emerald Ash Borer Project.* http://www.emeraldashborer.info/ (Consulté le March 2nd, 2020)

Engel P, Kwong WK, McFrederick Q, Anderson KE, Barribeau SM, Chandler JA, Cornman RS, Dainat J, De Miranda JR, Doublet V (2016) The bee microbiome: impact on bee health and model for evolution and ecology of host-microbe interactions. *MBio* 7(2).

Engel P & Moran NA (2013) The gut microbiota of insects–diversity in structure and function. *FEMS microbiology reviews* 37(5):699-735.

Eyles A, Bonello P, Ganley R, Mohammed C (2010) Induced resistance to pests and pathogens in trees. *New Phytologist* 185(4):893-908.

Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biological conservation* 61(1):1-10.

Faust K & Raes J (2012) Microbial interactions: from networks to models. *Nature Reviews Microbiology* 10(8):538-550.

Feldhaar H (2011) Bacterial symbionts as mediators of ecologically important traits of insect hosts. *Ecological Entomology* 36(5):533-543.

Ferguson LV, Dhakal P, Lebenzon JE, Heinrichs DE, Bucking C, Sinclair BJ (2018) Seasonal shifts in the insect gut microbiome are concurrent with changes in cold tolerance and immunity. *Functional ecology* 32(10):2357-2368.

Fischer R, Ostafe R, Twyman RM (2013) *Yellow Biotechnology II:Insect Biotechnology in Plant Protection and Industry,* Springer. p 51-64.

Fox-Dobbs K, Doak DF, Brody AK, Palmer TM (2010) Termites create spatial structure and govern ecosystem function by affecting N2 fixation in an East African savanna. *Ecology* 91(5):1296-1307.

Francese JA, Fraser I, Lance DR, Mastro VC (2011) Efficacy of multifunnel traps for capturing emerald ash borer (Coleoptera: Buprestidae): effect of color, glue, and other trap coatings. *Journal of economic entomology* 104(3):901-908.

Franzini PZ, Ramond J-B, Scholtz CH, Sole CL, Ronca S, Cowan DA (2016) The gut microbiomes of two Pachysoma MacLeay desert dung beetle species (Coleoptera: Scarabaeidae: Scarabaeinae) feeding on different diets. *PloS one* 11(8):e0161118.

Freeman LC, Roeder D, Mulholland RR (1979) Centrality in social networks: II. Experimental results. *Social networks* 2(2):119-141.

Freilich MA, Wieters E, Broitman BR, Marquet PA, Navarrete SA (2018) Species co-occurrence networks: Can they reveal trophic and non-trophic interactions in ecological communities? *Ecology* 99(3):690-699.

Gandhi KJ & Herms DA (2010a) North American arthropods at risk due to widespread Fraxinus mortality caused by the alien emerald ash borer. *Biological Invasions* 12(6):1839-1846.

Gandhi KJ & Herms DA (2010b) Direct and indirect effects of alien insect herbivores on ecological processes and interactions in forests of eastern North America. *Biological Invasions* 12(2):389-405.

García AMn, Moumen A, Ruiz DY, Alcaide EM (2003) Chemical composition and nutrients availability for goats and sheep of two-stage olive cake and olive leaves. *Animal Feed Science and Technology* 107(1-4):61-74.

Garg B (2012) *Plant Analysis: Comprehensive Methods and Protocols.* Scientific Publishers,

Gilbert SF, Sapp J, Tauber AI (2012) A symbiotic view of life: we have never been individuals. *The Quarterly review of biology* 87(4):325-341.

Glassman SI & Martiny JBH (2018) Broadscale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. *mSphere* 3(4):e00148-00118.

Gong X, Chen T-W, Zieger SL, Bluhm C, Heidemann K, Schaefer I, Maraun M, Liu M, Scheu S (2018) Phylogenetic and trophic determinants of gut microbiota in soil oribatid mites. *Soil Biology and Biochemistry* 123:155-164.

González AMM, Dalsgaard B, Olesen JM (2010) Centrality measures and the importance of generalist species in pollination networks. *Ecological complexity* 7(1):36-43.

Greenblum S, Turnbaugh PJ, Borenstein E (2012) Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proceedings of the National Academy of Sciences* 109(2):594-599.

Griffith DM, Veech JA, Marsh CJ (2016) Cooccur: probabilistic species co-occurrence analysis in R. *Journal of Statistical Software* 69(2):1-17.

Gupta A & Nair S (2020) Dynamics of Insect–Microbiome Interaction Influence Host and Microbial Symbiont. *Frontiers in microbiology* 11:1357.

Haack RA, Jendek E, Liu H, Marchant KR, Petrice TR, Poland TM, Ye H (2002) The Emerald Ash Borer: A New Exotic Pest in North America.

Hagerman AE (1988) Extraction of tannin from fresh and preserved leaves. *Journal of chemical Ecology* 14(2):453-461.

Harborne A (1998) *Phytochemical methods a guide to modern techniques of plant analysis.* springer science & business media,

Harborne JB (1988) *Introduction to Ecological Biochemistry.* Academic Press, New-York, 3

Hehemann J-H, Arevalo P, Datta MS, Yu X, Corzett CH, Henschel A, Preheim SP, Timberlake S, Alm EJ, Polz MF (2016) Adaptive radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes. *Nature communications* 7(1):1-10.

Henry LM, Peccoud J, Simon J-C, Hadfield JD, Maiden MJ, Ferrari J, Godfray HCJ (2013) Horizontally transmitted symbionts and host colonization of ecological niches. *Current biology* 23(17):1713-1717.

Herms DA & McCullough DG (2014) Emerald ash borer invasion of North America: history, biology, ecology, impacts, and management. *Annu Rev Entomol* 59:13-30.

Herms DA, Stone AK, Chatfield JA (2004) Emerald ash borer: the beginning of the end of Ash in North America? *Special Circular-Ohio Agricultural Research and Development Center* :62-71.

Hernández-García JA, Briones-Roblero CI, Rivera-Orduña FN, Zúñiga G (2017) Revealing the gut bacteriome of Dendroctonus bark beetles (Curculionidae: Scolytinae): diversity, core members and co-evolutionary patterns. *Scientific Reports* 7(1):13864.

Horner-Devine MC & Bohannan BJ (2006) Phylogenetic clustering and overdispersion in bacterial communities. *Ecology* 87(sp7):S100-S108.

Howe GA & Schaller A (2008) Direct defenses in plants and their induction by wounding and insect herbivores. *Induced plant resistance to herbivory,* Springer. p 7-29.

Hsieh T, Ma K, Chao A (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods in Ecology and Evolution* 7(12):1451-1456.

Hu X, Yu J, Wang C, Chen H (2014) Cellulolytic bacteria associated with the gut of Dendroctonus armandi larvae (Coleoptera: Curculionidae: Scolytinae). *Forests* 5(3):455-465.

Huang S & Zhang H (2013) The impact of environmental heterogeneity and life stage on the hindgut microbiota of Holotrichia parallela larvae (Coleoptera: Scarabaeidae). *PLoS One* 8(2):e57169.

Hughes CC, Johns RC, Sweeney JD (2014) A technical guide to installing beetle traps in the upper crown of trees. *Journal of the Acadian Entomological Society* 10.

Humphrey PT & Whiteman NK (2020) Insect herbivory reshapes a native leaf microbiome. *Nature Ecology & Evolution* 4(2):221-229.

Hunter PJ, Hand P, Pink D, Whipps JM, Bending GD (2010) Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (Lactuca species) phyllosphere. *Applied and environmental microbiology* 76(24):8117-8125.

Hurst GDD (2017) Extended genomes: symbiosis and evolution. *Interface Focus* 7(5):20170001.

Itoh H, Jang S, Takeshita K, Ohbayashi T, Ohnishi N, Meng X-Y, Mitani Y, Kikuchi Y (2019) Host–symbiont specificity determined by microbe–microbe competition in an insect gut. *Proceedings of the National Academy of Sciences* 116(45):22673-22682.

Jendek E & Grebennikov V (2011) *Agrilus (Coleoptera, Buprestidae) of East Asia.* Nakladatelstvi Jan Farkač,

Jordan HR & Tomberlin JK (2017) Abiotic and biotic factors regulating inter-kingdom engagement between insects and microbe activity on vertebrate remains. *Insects* 8(2):54.

Kadivar H & Stapleton AE (2003) Ultraviolet radiation alters maize phyllosphere bacterial diversity. *Microbial Ecology* 45(4):353-361.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26(11):1463-1464.

Kembel SW, Eisen JA, Pollard KS, Green JL (2011) The phylogenetic diversity of metagenomes. *PLoS One* 6(8):e23214.

Kembel SW & Hubbell SP (2006) The phylogenetic structure of a neotropical forest tree community. *Ecology* 87(sp7):S86-S99.

Kembel SW, O'Connor TK, Arnold HK, Hubbell SP, Wright SJ, Green JL (2014) Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proceedings of the National Academy of Sciences* 111(38):13715-13720.

Kim JM, Choi M-Y, Kim J-W, Lee SA, Ahn J-H, Song J, Kim S-H, Weon H-Y (2017) Effects of diet type, developmental stage, and gut compartment in the gut bacterial communities of two Cerambycidae species (Coleoptera). *Journal of Microbiology* 55(1):21-30.

Klooster WS, Gandhi KJ, Long LC, Perry KI, Rice KB, Herms DA (2018) Ecological impacts of emerald ash borer in forests at the epicenter of the invasion in North America. *Forests* 9(5):250.

Klooster WS, Herms DA, Knight KS, Herms CP, McCullough DG, Smith A, Gandhi KJ, Cardina J (2014) Ash (Fraxinus spp.) mortality, regeneration, and seed bank dynamics in mixed hardwood forests following invasion by emerald ash borer (Agrilus planipennis). *Biological Invasions* 16(4):859-873.

Knief C (2014) Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Frontiers in plant science* 5.

Knight KS, Brown JP, Long RP (2013) Factors affecting the survival of ash (Fraxinus spp.) trees infested by emerald ash borer (Agrilus planipennis). *Biological Invasions* 15(2):371-383.

Kostova I & Iossifova T (2007) Chemical components of Fraxinus species. *Fitoterapia* 78(2):85-106.

Kovacs KF, Haight RG, McCullough DG, Mercader RJ, Siegert NW, Liebhold AM (2010) Cost of potential emerald ash borer damage in U.S. communities, 2009–2019. *Ecological Economics* 69(3):569-578.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* 33(7):1870-1874.

Lahti L, Shetty S, Salojarvi J (2017) Tools for microbiome analysis in R. .

Layeghifard M, Hwang DM, Guttman DS (2017) Disentangling interactions in the microbiome: a network perspective. *Trends in microbiology* 25(3):217-228.

Legendre P & De Cáceres M (2013) Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology letters* 16(8):951-963.

Legendre P & Legendre LF (2012) *Numerical ecology.* Elsevier,

Lennon JT, Muscarella ME, Placella SA, Lehmkuhl BK (2018) How, When, and Where Relic DNA Affects Microbial Diversity. *mBio* 9(3):e00637-00618.

Lieutier F, Yart A, Salle A (2009) Stimulation of tree defenses by Ophiostomatoid fungi can explain attack success of bark beetles on conifers. *Annals of Forest Science* 66(8):801-801.

Liu H, Bauer LS, Gao R, Zhao T, Petrice TR, Haack RA (2004) Exploratory survey for the emerald ash borer, Agrilus planipennis (Coleoptera: Buprestidae), and its natural enemies in China. *Great Lakes Entomologist* 36(3-4)(Fall-Winter 2003).

Lyons DB, Caister C, De Groot P, Hamilton B, Marchant K, Scarr T, Turgeon J (2007) Survey guide for detection of emerald ash borer.

Lyons DB, Lavallée R, Kyei–Poku G, Van Frankenhuyzen K, Johny S, Guertin C, Francese JA, Jones GC, Blais M (2012) Towards the Development of an Autocontamination Trap System to Manage Populations of Emerald Ash Borer (Coleoptera: Buprestidae) With the

Native Entomopathogenic Fungus, Beauveria bassiana. *Journal of Economic Entomology* 105(6):1929-1939.

MacFarlane DW & Meyer SP (2005) Characteristics and distribution of potential ash tree hosts for emerald ash borer. *Forest Ecology and Management* 213(1):15-24.

Margulis L (1992) Biodiversity: molecular biological domains, symbiosis and kingdom origins. *Biosystems* 27(1):39-51.

Margulis L (1993) Origins of species: acquired genomes and individuality. *Biosystems* 31(2-3):121-125.

Mariño YA, Ospina OE, Verle Rodrigues JC, Bayman P (2018) High diversity and variability in the bacterial microbiota of the coffee berry borer (Coleoptera: Curculionidae), with emphasis on Wolbachia. *Journal of Applied Microbiology* 125(2):528-543.

Marquis RJ (2004) Herbivores rule. *Science* 305(5684):619-621.

Marquis RJ, Newell EA, Villegas AC (1997) Non-structural carbohydrate accumulation and use in an understorey rain-forest shrub and relevance for the impact of leaf herbivory. *Functional Ecology* 11(5):636-643.

Mascarin GM & Jaronski ST (2016) The production and uses of Beauveria bassiana as a microbial insecticide. *World Journal of Microbiology and Biotechnology* 32(11):177.

Mason CJ, Hanshew AS, Raffa KF (2015) Contributions by Host Trees and Insect Activity to Bacterial Communities in Dendroctonus valens (Coleoptera: Curculionidae) Galleries, and Their High Overlap With Other Microbial Assemblages of Bark Beetles. *Environmental Entomology* 45(2):348-356.

McCullough DG (2020) Challenges, tactics and integrated management of emerald ash borer in North America. *Forestry: An International Journal of Forest Research* 93(2):197-211.

McCullough DG, Poland TM, Tluczek AR, Anulewicz A, Wieferich J, Siegert NW (2019) Emerald ash borer (Coleoptera: Buprestidae) densities over a 6-yr period on untreated trees and trees treated with systemic insecticides at 1-, 2-, and 3-yr intervals in a Central Michigan Forest. *Journal of Economic Entomology* 112(1):201-212.

McCullough DG, Siegert NW, Cappaert D, Poland TM, McDonald R (2007) Sinks, bark, and garlon: applied studies for emerald ash borer management. *Emerald ash borer and asian longhorhed beetle research and development review*. p 92.

Medina R, Nachappa P, Tamborindeguy C (2011) Differences in bacterial diversity of host-associated populations of Phylloxera notabilis Pergande (Hemiptera: Phylloxeridae) in pecan and water hickory. *Journal of Evolutionary Biology* 24(4):761-771.

Mittapalli O, Bai X, Mamidala P, Rajarapu SP, Bonello P, Herms DA (2010) Tissue-specific transcriptomics of the exotic invasive insect pest emerald ash borer (Agrilus planipennis). *PLoS one* 5(10):e13708.

Mogouong J, Constant P, Lavallée R, Guertin C (2020) Gut microbiome of the emerald ash borer, Agrilus planipennis Fairmaire, and its relationship with insect population density. *FEMS Microbiology Ecology* 10.1093/femsec/fiaa141.

Mogouong J, Constant P, Legendre P, Guertin C (2021) The phyllosphere microbiome of host trees contributes more than leaf phytochemicals to variation in the Agrilus planipennis Fairmaire gut microbiome structure. *Scientific Reports* 11(1):1-15.

Montagna M, Chouaia B, Mazza G, Prosdocimi EM, Crotti E, Mereghetti V, Vacchini V, Giorgi A, De Biase A, Longo S (2015a) Effects of the diet on the microbiota of the red palm weevil (Coleoptera: Dryophthoridae). *PLoS One* 10(1):e0117439.

Montagna M, Gómez-Zurita J, Giorgi A, Epis S, Lozzia G, Bandi C (2015b) Metamicrobiomics in herbivore beetles of the genus Cryptocephalus (Chrysomelidae): toward the understanding of ecological determinants in insect symbiosis. *Insect Science* 22(3):340-352.

Montoya D, Yallop M, Memmott J (2015) Functional group diversity increases with modularity in complex food webs. *Nature communications* 6(1):1-9.

Morales-Jiménez J, Zúñiga G, Ramírez-Saad HC, Hernández-Rodríguez C (2012) Gut-associated bacteria throughout the life cycle of the bark beetle Dendroctonus rhizophagus Thomas and Bright (Curculionidae: Scolytinae) and their cellulolytic activities. *Microbial Ecology* 64(1):268-278.

Morales-Jiménez J, Zúñiga G, Villa-Tanaca L, Hernández-Rodríguez C (2009) Bacterial community and nitrogen fixation in the red turpentine beetle, Dendroctonus valens LeConte (Coleoptera: Curculionidae: Scolytinae). *Microbial Ecology* 58(4):879-891.

Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. *Annual review of genetics* 42:165-190.

Moran NA & Yun Y (2015) Experimental replacement of an obligate insect symbiont. *Proceedings of the National Academy of Sciences* 112(7):2093-2096.

Mrazek J, Strosova L, Fliegerova K, Kott T, Kopecny J (2008) Diversity of insect intestinal microflora. *Folia Microbiol (Praha)* 53(3):229-233.

Nadarasah G & Stavrinides J (2014) Quantitative evaluation of the host-colonizing capabilities of the enteric bacterium Pantoea using plant and insect hosts. *Microbiology* 160(3):602-615.

Nardi JB, Mackie RI, Dawson JO (2002) Could microbial symbionts of arthropod guts contribute significantly to nitrogen fixation in terrestrial ecosystems? *J Insect Physiol* 48(8):751-763.

Newman ME & Girvan M (2004) Finding and evaluating community structure in networks. *Physical review E* 69(2):026113.

Nikoh N, Hosokawa T, Oshima K, Hattori M, Fukatsu T (2011) Reductive evolution of bacterial genome in insect gut environment. *Genome Biol Evol* 3:702-714.

Oksanen J, Blanchet F, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin P, O'Hara R, Simpson G, Solymos P (2019) Package 'vegan'. R package version 2.5-6.

Oksanen J, Kindt R, Legendre P, O'Hara B, Stevens MHH, Oksanen MJ, Suggests M (2007) The vegan package. *Community ecology package* 10:631-637.

Olesen JM, Bascompte J, Dupont YL, Jordano P (2007) The modularity of pollination networks. *Proceedings of the National Academy of Sciences* 104(50):19891-19896.

Oliverio AM, Gan H, Wickings K, Fierer N (2018) A DNA metabarcoding approach to characterize soil arthropod communities. *Soil Biology and Biochemistry* 125:37-43.

Orlova-Bienkowskaja MJ (2014) Ashes in Europe are in danger: the invasive range of Agrilus planipennis in European Russia is expanding. *Biological Invasions* 16(7):1345-1349.

Orlova-Bienkowskaja MJ, Drogvalenko AN, Zabaluev IA, Sazhnev AS, Peregudova EY, Mazurov SG, Komarov EV, Struchaev VV, Martynov VV, Nikulina TV (2020) Current range of

Agrilus planipennis Fairmaire, an alien pest of ash trees, in European Russia and Ukraine. *Annals of Forest Science* 77(2):1-14.

Orlova-Bienkowskaja MJ & Bieńkowski AO (2016) The life cycle of the emerald ash borer Agrilus planipennis in European Russia and comparisons with its life cycles in Asia and North America. *Agricultural and Forest Entomology* 18(2):182-188.

Panteleev DY, Goryacheva II, Andrianov BV, Reznik NL, Lazebny OE, Kulikov AM (2007) The endosymbiotic bacterium Wolbachia enhances the nonspecific resistance to insect pathogens and alters behavior of Drosophila melanogaster. *Russian Journal of Genetics* 43(9):1066-1069.

Paracer S & Ahmadjian VS (1999) *Symbiosis : an introduction to biological associations.* Oxford University Press, New York, 2nd ed.

Paré PW & Tumlinson JH (1999) Plant volatiles as a defense against insect herbivores. *Plant physiology* 121(2):325-332.

Peres-Neto PR, Legendre P, Dray S, Borcard D (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* 87(10):2614-2625.

Poland TM & McCullough DG (2006) Emerald ash borer: invasion of the urban forest and the threat to North America's ash resource. *Journal of Forestry* 104(3):118-124.

Poland TM, McCullough DG, dr Groot P, Grant G, MacDonald L, Cappaert DL (2005) Progress toward developing trapping techniques for the emerald ash borer. *In: Mastro, Victor; Reardon, Richard, comps. Proceedings of the emerals ash borer research and development meeting; 2004 October 5-6; Romulus, MI. FHTET 2004-15. Morgantown, WV: US Forest Service, Forest Health Technology Enterprise Team: 53-54.*

Popa V, Deziel E, Lavallee R, Bauce E, Guertin C (2012) The complex symbiotic relationships of bark beetles with microorganisms: a potential practical approach for biological control in forestry. *Pest Manag Sci* 68(7):963-975.

Priya NG, Ojha A, Kajla MK, Raj A, Rajagopal R (2012) Host plant induced variation in gut bacteria of Helicoverpa armigera. *PloS one* 7(1):e30768.

Pureswaran DS & Poland TM (2009) Host selection and feeding preference of Agrilus planipennis (Coleoptera: Buprestidae) on ash (Fraxinus spp.). *Environmental Entomology* 38(3):757-765.

Qadri M, Short S, Gast K, Hernandez J, Wong AC-N (2020) Microbiome Innovation in Agriculture: Development of Microbial Based Tools for Insect Pest Management. *Front. Sustain. Food Syst. 4: 547751. doi: 10.3389/fsufs*.

Rani A, Sharma A, Rajagopal R, Adak T, Bhatnagar RK (2009) Bacterial diversity analysis of larvae and adult midgut microflora using culture-dependent and culture-independent methods in lab-reared and field-collected Anopheles stephensi-an Asian malarial vector. *BMC microbiology* 9(1):96.

Rodriguez-Saona C, Poland TM, Miller JR, Stelinski LL, Grant GG, Groot Pd, Buchan L, MacDonald L (2006) Behavioral and electrophysiological responses of the emerald ash borer, Agrilus planipennis, to induced volatiles of Manchurian ash, Fraxinus mandshurica. *CHEMOECOLOGY* 16(2):75-86.

Roe AD, Rice AV, Coltman DW, Cooke JEK, Sperling FAH (2011) Comparative phylogeography, genetic differentiation and contrasting reproductive modes in three fungal symbionts of a multipartite bark beetle symbiosis. *Molecular Ecology* 20(3):584-600.

Rosenthal GA & Berenbaum MR (2012) *Herbivores: their interactions with secondary plant metabolites: ecological and evolutionary processes.* Academic Press,

Roy S & Banerjee D (2018) Diversity of Endophytes in Tropical Forests. *Endophytes of Forest Trees,* Springer. p 43-62.

Ryall KL, Fidgen JG, Turgeon JJ (2011) Detectability of the emerald ash borer (Coleoptera: Buprestidae) in asymptomatic urban trees by using branch samples. *Environmental Entomology* 40(3):679-688.

Santos-Garcia D, Mestre-Rincon N, Zchori-Fein E, Morin S (2020) Inside out: microbiota dynamics during host-plant adaptation of whiteflies. *The ISME Journal* 14(3):847-856.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology* 75(23):7537-7541.

Schoonhoven LM, Van Loon B, van Loon JJ, Dicke M (2005) *Insect-plant biology.* Oxford University Press on Demand,

Schowalter TD (2016) *Insect ecology: an ecosystem approach.* Academic press,

Scott DW (2009) Sturges' rule. *Wiley Interdisciplinary Reviews: Computational Statistics* 1(3):303-306.

Shikano I, Rosa C, Tan C-W, Felton GW (2017) Tritrophic interactions: microbe-mediated plant effects on insect herbivores. *Annual review of Phytopathology* 55:313-331.

Siegert NW, McCullough D, Williams D, Fraser I, Poland T, Pierce S (2010) Dispersal of Agrilus planipennis (Coleoptera: Buprestidae) from discrete epicenters in two outlier sites. *Environmental Entomology* 39(2):253-265.

Siegert NW, McCullough DG, Liebhold AM, Telewski FW (2014) Dendrochronological reconstruction of the epicentre and early spread of emerald ash borer in N orth A merica. *Diversity and distributions* 20(7):847-858.

Simon J-C, Marchesi JR, Mougel C, Selosse M-A (2019) Host-microbiota interactions: from holobiont theory to analysis. *Microbiome* 7(1):1-5.

Slansky F & Haack R (1986) Age-specific flight behavior in relation to body weight and lipid content of Ips calligraphus reared in slash pine bolts with thick or thin inner bark (phloem). *Entomologia Experimentalis et Applicata* 40(2):197-207.

Slesak RA, Lenhart CF, Brooks KN, D'Amato AW, Palik BJ (2014) Water table response to harvesting and simulated emerald ash borer mortality in black ash wetlands in Minnesota, USA. *Canadian Journal of Forest Research* 44(8):961-968.

Smith CC, Snowberg LK, Caporaso JG, Knight R, Bolnick DI (2015) Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *The ISME journal* 9(11):2515-2526.

Snyder RL (1985) Hand calculating degree days. *Agricultural and forest meteorology* 35(1-4):353-358.

Soul L & Wright D (2020) Phylogenetic Comparative Methods: A User's Guide for Paleontologists.

Srei N, Guertin C, Lavallée R, Lajoie M-È, Brousseau C, Bergevin R, Miller F, McMillin K, Trudel R (2020a) Microbial Control of the Emerald Ash Borer (Coleoptera: Buprestidae) Using

Beauveria bassiana (Hypocreales: Cordycipitaceae) by the Means of an Autodissemination Device. *Journal of Economic Entomology* 113(6):2657-2665.

Srei N, Lavallée R, Guertin C (2020b) Horizontal transmission of the entomopathogenic fungal isolate INRS-242 of Beauveria bassiana (Hypocreales: Cordycipitaceae) in Emerald Ash Borer, Agrilus planipennis (Coleoptera: Buprestidae). *Journal of Economic Entomology* 113(1):543-545.

Stam JM, Kroes A, Li Y, Gols R, van Loon JJ, Poelman EH, Dicke M (2014) Plant interactions with multiple insect herbivores: from community to genes. *Annual review of plant biology* 65:689-713.

Starnes RL, Liu CL, Marrone PG (1993) History, use, and future of microbial insecticides. *American Entomologist* 39(2):83-91.

Steinhaus EA (1958) Crowding as a possible stress factor in insect disease. *Ecology* 39(3):503-514.

Team RC (2013) R: A language and environment for statistical computing.

Thapa S & Prasanna R (2018) Prospecting the characteristics and significance of the phyllosphere microbiome. *Annals of microbiology* 68(5):229-245.

Toju H, Tanabe AS, Yamamoto S, Sato H (2012) High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PloS one* 7(7):e40863.

Torti SD, Dearing MD, Kursar TA (1995) Extraction of phenolic compounds from fresh leaves: a comparison of methods. *Journal of Chemical Ecology* 21(2):117-125.

USDA (2009) Emerald Ash Borer. *Michigan State University Michigan USA*.

Vacheron J, Péchy-Tarr M, Brochet S, Heiman CM, Stojiljkovic M, Maurhofer M, Keel C (2019) T6SS contributes to gut microbiome invasion and killing of an herbivorous pest insect by plant-beneficial Pseudomonas protegens. *The ISME Journal* 13(5):1318-1329.

Valenta V, Moser D, Kapeller S, Essl F (2017) A new forest pest in Europe: a review of Emerald ash borer (Agrilus planipennis) invasion. *Journal of Applied Entomology* 141(7):507-526.

Vamosi JC & Vamosi SM (2007) Body size, rarity, and phylogenetic community structure: insights from diving beetle assemblages of Alberta. *Diversity and Distributions* 13(1):1-10.

van der Heijden MG & Hartmann M (2016) Networking in the plant microbiome. *PLoS Biology* 14(2):e1002378.

Van Soest Pv, Robertson J, Lewis B (1991) Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *Journal of dairy science* 74(10):3583-3597.

van Veelen HPJ, Salles JF, Tieleman BI (2017) Multi-level comparisons of cloacal, skin, feather and nest-associated microbiota suggest considerable influence of horizontal acquisition on the microbiota assembly of sympatric woodlarks and skylarks. *Microbiome* 5(1):1-17.

Vasanthakumar A, Handelsman J, Schloss PD, Bauer LS, Raffa KF (2008) Gut microbiota of an invasive subcortical beetle, Agrilus planipennis Fairmaire, across various life stages. *Environmental Entomology* 37(5):1344-1353.

Veech JA (2013) A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography* 22(2):252-260.

Venables W, Ripley B, Isbn S (2002) Statistics Complements to Modern Applied Statistics with S Fourth edition by.

Villari C, Herms DA, Whitehill JG, Cipollini D, Bonello P (2016) Progress and gaps in understanding mechanisms of ash tree resistance to emerald ash borer, a model for wood-boring insects that kill angiosperms. *New Phytologist* 209(1):63-79.

Villari C, Whitehill JG, Cipollini D, Herms DA, Bonello P (2014) Mechanisms of Ash Resistance to Emerald Ash Borer: Progress and Gaps.

Vorholt JA (2012) Microbial life in the phyllosphere. *Nature Reviews Microbiology* 10(12):828.

Walter J, Britton RA, Roos S (2011) Host-microbial symbiosis in the vertebrate gastrointestinal tract and the Lactobacillus reuteri paradigm. *Proceedings of the National Academy of Sciences* 108(Supplement 1):4645-4652.

Walterson AM & Stavrinides J (2015) Pantoea: insights into a highly versatile and diverse genus within the Enterobacteriaceae. *FEMS Microbiology Reviews* 39(6):968-984.

Wang, Yang Z, Gould J, Zhang Y, Liu G, Liu E (2010) The biology and ecology of the emerald ash borer, Agrilus planipennis, in China. *J Insect Sci* 10:128.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology* 73(16):5261-5267.

Watanabe H & Tokuda G (2010) Cellulolytic systems in insects. *Annual review of entomology* 55.

Webb CO (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist* 156(2):145-155.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ (2002) Phylogenies and community ecology. *Annual review of ecology and systematics* 33(1):475-505.

Wei X, Reardon D, Wu Y, Sun J-H (2004) Emerald ash borer, Agrilus planipennis Fairmaire (Coleoptera: Buprestidae), in China: a review and distribution survey. *Acta Entomologica Sinica* 47(5):679-685.

Wei X, Wu Y, Reardon R, SUN TH, Lu M, SUN JH (2007) Biology and damage traits of emerald ash borer (*Agrilus planipennis* Fairmaire) in China. *Insect Science* 14(5):367-373.

Wellburn R (1994) The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *Journal of plant physiology* 144(3):307-313.

Whitaker D & Christman M (2014) Package 'Clustsig'R Package Version 1.1 2014. *Disponible online en: https://cran. rproject. org/web/packages/clustsig/clustsig. pdf*.

Wielkopolan B & Obrepalska-Steplowska A (2016) Three-way interaction among plants, bacteria, and coleopteran insects. *Planta* 244(2):313-332.

Williams RJ, Howe A, Hofmockel KS (2014) Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. *Frontiers in microbiology* 5:358.

Winans NJ, Walter A, Chouaia B, Chaston JM, Douglas AE, Newell PD (2017) A genomic investigation of ecological differentiation between free-living and Drosophila-associated bacteria. *Molecular ecology* 26(17):4536-4550.

Winder RS, Macey DE, Cortese J (2010) Dominant bacteria associated with broods of mountain pine beetle, Dendroctonus ponderosae (Coleoptera: Curculionidae, Scolytinae). *Journal of the Entomological Society of British Columbia* 107:43-56.

Wu J (2016) *Transcriptomics and Gene Regulation.* Springer,

Wylie FR & Speight MR (2012) *Insect pests in tropical forestry.* CABI, Oxfordshire, UK

Xu L, Lou Q, Cheng C, Lu M, Sun J (2015) Gut-associated bacteria of Dendroctonus valens and their involvement in verbenone production. *Microbial ecology* 70(4):1012-1023.

Youngblut ND, Reischer GH, Walters W, Schuster N, Walzer C, Stalder G, Ley RE, Farnleitner AH (2019) Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. *Nature communications* 10(1):1-15.

Yun J-H, Roh SW, Whon TW, Jung M-J, Kim M-S, Park D-S, Yoon C, Nam Y-D, Kim Y-J, Choi J-H (2014) Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Applied and environmental microbiology* 80(17):5254-5264.

Yutthammo C, Thongthammachat N, Pinphanichakarn P, Luepromchai E (2010) Diversity and activity of PAH-degrading bacteria in the phyllosphere of ornamental plants. *Microbial ecology* 59(2):357-368.

Zhang Z, Jiao S, Li X, Li M (2018) Bacterial and fungal gut communities of Agrilus mali at different developmental stages and fed different diets. *Scientific Reports* 8(1):15634.

Zhao L, Lu M, Niu H, Fang G, Zhang S, Sun J (2013) A native fungal symbiont facilitates the prevalence and development of an invasive pathogen–native vector symbiosis. *Ecology* 94(12):2817-2826.

Zilber-Rosenberg I & Rosenberg E (2008) Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS microbiology reviews* 32(5):723-735.

Zindel R, Gottlieb Y, Aebi A (2011) Arthropod symbioses: a neglected parameter in pest- and disease-control programmes. *Journal of Applied Ecology* 48(4):864-872.