UNIVERSITÉ DU QUÉBEC À MONTRÉAL

FACTEURS DÉTERMINANT LA COMPOSITION ET LA CO-OCCURRENCE DU MICROBIOME DE LA TORDEUSE DES BOURGEONS DE L'ÉPINETTE

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DU DOCTORAT EN BIOLOGIE

PAR
MATHIEU LANDRY

FÉVRIER 2022

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.04-2020). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Tout d'abord, je tiens à remercier mon directeur, le Dr. Steven W. Kembel, pour sa grande compréhension, sa constante disponibilité et surtout son souci du succès de ses étudiants. C'est grâce à son support ininterrompu que j'ai pu acquérir les notions en écologie qui m'étaient complètement étrangères au début de mon doctorat et qui m'étaient nécessaires à son accomplissment. Je remercie aussi mon co-directeur, le Dr. Patrick M.A. James, dont l'apport en termes de conseils de rédaction, de statistiques et d'écologie a grandement contribué à la qualité de cette thèse. Je remercie aussi le Dr. Daniel Kneeshaw, qui a toujours su poser les bonnes questions pour améliorer le projet. Je remercie Dr. Rob Johns, Dr. Paul Whalen, Jérémie Brisson et Marie-Line Buteau pour leur contribution importante à la collecte d'échantillons. Je remercie Geneviève Bourret sans laquelle je n'aurais jamais réussi à amplifier mes échantillons. Au travers du long processus qu'est le programme de doctorat, j'ai eu la chance de bénéficier du support émotionnel de mes proches, parmi lesquels je veux souligner Béatrice Bérubé, Louis-Marie Landry et Martine Lebrun. Partager les haut et bas du programme avec mes collègues Mona Parizadeh et Pedro Henrique Pereira Braga fut aussi d'une grande aide pour garder le moral tout au long du processus.

DÉDICACE

The universe is under no obligation to make sense to you.

Neil deGrasse Tyson

AVANT-PROPOS

Cette thèse consiste en trois articles scientifiques rédigés en anglais. J'ai été responsable du désign expérimental, de l'échantillonnage, du travail en laboratoire, de l'analyse bioinformatique et statistique, et de la rédaction et révision des articles. Steven W. Kembel, Patrick M.A James et Daniel Kneeshaw ont tous participé dans le design expérimental et la révision des trois articles. Je suis le premier auteur des trois articles.

Le premier article, *Spruce budworm bacterial communities vary among sites and host tree species in a boreal landscape*, a été publié par le journal *Journal of Biogeography*. Les co-auteurs sont Steven W. Kembel, Patrick M.A James et Daniel Kneeshaw.

Le deuxième article, *Eastern spruce budworm gut communities differ between sexes and between gut tissues and whole larvae*, est en préparation. Les co-auteurs sont Steven W. Kembel, Patrick M.A James et Daniel Kneeshaw.

Le troisième article, *Variation in spruce budworm microbial association networks among whole larvae and larval gut tissues, host tree species and larval sex*, est en préparation. Les co-auteurs sont Steven W. Kembel, Patrick M.A James et Daniel Kneeshaw.

TABLE DES MATIÈRES

AVANT-PROPOS iv

LISTE DES FIGURES x

LISTE DES TABLEAUX xv

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES xvi

RÉSUMÉ xvii

ABSTRACT xix

INTRODUCTION 1

0.1 Définition du microbiome 1

0.2 Les microbiomes d'insectes 2

0.3 La tordeuse des bourgeons de l'épinette et son impact 3

0.4 Les microbiomes de Lépidoptères 5

0.5 Pourquoi étudier le microbiome de la TBE 7

0.6 Processus dirigeant la formation de communautés microbiennes 7

0.7 Les facteurs de variation spatiale et les « Moran eigenvectors maps » 9

0.8 Les facteurs environnementaux 10

0.9 Les différents arbres-hôtes 10

0.10 Différences du microbiome reliées au sexe de l'hôte 11

0.11 Variation du microbiome à travers l'hôte 12

0.12 Les réseaux de co-occurrence microbiens 13

0.13 Présentation de la thèse 14

0.13.1 Chapitre 1 15

0.13.2 Chapitre 2 16

0.13.3 Chapitre 3 17

CHAPITRE I SPRUCE BUDWORM BACTERIAL COMMUNITIES VARY AMONG SITES AND HOST TREE SPECIES IN A BOREAL LANDSCAPE..... 20

1.1 Résumé 21

1.2 Abstract..... 22

1.3 Introduction 23

1.4 Methods 26

1.4.1 Study area 26
1.4.2 Sample collection 28
1.4.3 Processing..... 28
1.4.4 Statistical analyses..... 29
1.4.5 Characterizing spatial structure 29
1.4.6 Larval and foliar communitiy variation..... 30
1.4.7 Correlation between foliar and larval bacterial communities..... 30
1.4.8 Modeling variation in bacterial community composition..... 31

1.5 Results 32

1.5.1 Larval and foliar community variation..... 32
1.5.2 Modeling variation in bacterial community structure 36
1.5.3 Correlation between foliar and larval bacterial communities..... 39

1.6 Discussion..... 40

1.7 Conclusions 44

1.8 Acknowledgements 45

CHAPITRE II EASTERN SPRUCE BUDWORM BACTERIAL COMMUNITIES DIFFER BETWEEN SEXES AND BETWEEN GUT TISSUES AND WHOLE LARVAE 46

2.1 Résumé 47

2.2 Abstract..... 48

2.3 Introduction 49

2.4 Methods 52

2.4.1 Study area 52
2.4.2 Sample collection 53
2.4.3 Sequencing and sequence data processing 54
2.4.4 Removal of contaminant sequences 54

2.4.5 Statistical analyses 55
2.5 Results 56
2.5.1 Larval sex differences 56
2.5.2 Bacterial communities in gut tissues versus whole larvae 63
2.6 Discussion 68
2.6.1 Larval sex differences 68
2.6.2 Bacterial communities in gut tissues versus whole larvae 69
2.7 Conclusions 73
2.8 Acknowledgements 73

CHAPITRE III VARIATION IN SPRUCE BUDWORM BACTERIAL ASSOCIATION NETWORKS AMONG WHOLE LARVAE AND LARVAL GUT TISSUES, HOST TREE SPECIES AND LARVAL SEX 74
3.1 Résumé 75
3.2 Abstract 76
3.3 Introduction 77
3.4 Methods 82
3.4.1 Whole larvae dataset 82
3.4.2 Gut tissues dataset 83
3.4.3 Sequencing and sequence data processing 83
3.4.4 Removal of contaminant sequences 84
3.4.5 Network inference 84
3.5 Results 85
3.5.1 Whole larvae and larval gut tissues 85
3.5.2 Host tree species variation 91
3.5.3 Larval sex variation 95
3.6 Discussion 98
3.6.1 Whole larvae and larval gut tissues 99
3.6.2 Host tree species variation 100
3.6.3 Larval sex variation 101
3.7 Conclusions 102
3.8 Acknowledgements 103

CONCLUSION 104

4.1 Principaux facteurs expliquant la variation spatiale du microbiome de la TBE 105
4.2 Variation entre les communautés bactériennes de la TBE en fonction du sexe et entre les communautés d'intestins et de larves entières 106
4.3 Les co-occurrences des communautés bactériennes de la TBE et leur variation 107
4.4 Limites du projet 108
4.5 Recommandations pour le futur 111

ANNEXE A FIGURE SUPPLÉMENTAIRE 1.1 115

ANNEXE B FIGURE SUPPLÉMENTAIRE 1.2 116

ANNEXE C FIGURE SUPPLÉMENTAIRE 1.3 117

ANNEXE D TABLEAU SUPPLÉMENTAIRE 1.1 118

ANNEXE E TABLEAU SUPPLÉMENTAIRE 1.2 119

ANNEXE F TABLEAU SUPPLÉMENTAIRE 1.3 120

ANNEXE G TABLEAU SUPPLÉMENTAIRE 1.4 121

ANNEXE H TABLEAU SUPPLÉMENTAIRE 1.5 122

ANNEXE I TABLEAU SUPPLÉMENTAIRE 1.6 123

ANNEXE J FIGURE SUPPLÉMENTAIRE 2.1 124

ANNEXE K FIGURE SUPPLÉMENTAIRE 2.2 125

ANNEXE L TABLEAU SUPPLÉMENTAIRE 2.1 126

ANNEXE M TABLEAU SUPPLÉMENTAIRE 2.2 127

ANNEXE N TABLEAU SUPPLÉMENTAIRE 2.3 128

ANNEXE O Figure SUPPLÉMENTAIRE 3.1 130

ANNEXE P Figure SUPPLÉMENTAIRE 3.2 131

RÉFÉRENCES..... 132

LISTE DES FIGURES

Figure 1.1: A) Map of sampling sites in Québec, Canada and surrounding provinces. Projection: NAD 1983 Québec Lambert Conformal. Sources: ESRI Basemap, MERN 2010. Author: CFR, 2021. B) Schematic of the sampling design at each site. 27

Figure 1.2: ANCOM-BC discriminant analysis of differentially abundant ASVs in spruce budworm larvae coming from different host tree species. Most differentially abundant ASVs were positively associated with balsam fir and were from the *Proteobacteria* and *Actinobacteria* phyla. 34

Figure 1.3: Variance partitioning analysis dividing variation in spruce budworm bacterial community structure among four possible predictors: environmental variables, foliage bacterial communities, host tree species and spatial structure (Moran's eigenvector maps (MEMs)). Values in the figure represent the proportion of variation in bacterial community structure explained by each set of variables and were determined based on a partial redundancy analysis of bacterial community composition. 35

Figure 1.4: Principal coordinates (PCoA) ordination based on the Bray-Curtis dissimilarity between bacterial communities, Vectors indicate variables that were significantly correlated with ordination axes ($p < 0.05$) based on a randomization test. Variance explained by each PCoA axis is shown in axis labels. Most predictor variables clustered into two separate groups of intercorrelated vectors that were correlated with the two ordination axes, indicating that the composition of spruce budworm bacterial communities covaried with several measured environmental variables. 38

Figure 1.5: ANCOM-BC discriminant analysis of differentially abundant ASVs between larvae and foliage they were feeding on. Most differentially abundant ASVs were from the *Proteobacteria* and *Firmicutes* phyla. 40

Figure 2.1: Map of sample collection sites in Quebec, Canada. 53

Figure 2.2: Mean relative abundance of all bacterial phyla detected in spruce budworm bacterial communities of gut tissues and whole larvae. 58

Figure 2.3: NMDS ordinations of spruce budworm gut tissue bacterial communities based on Bray-Curtis dissimilarity for each site, female gut tissue communities are shown in red and male gut tissue communities in blue. 68% confidence interval ellipses are shown overlaid on the ordination for each group. 59

Figure 2.4: Boxplots showing comparisons between bacterial communities of gut tissue samples of male and female spruce budworm larvae in terms of Shannon index of bacterial communities. P-values of Mann-Whitney-Wilcoxon tests between male and female community Shannon index of individual gut tissues are shown over each boxplot. 60

Figure 2.5: DeSEQ2 discriminant analysis of differentially abundant bacterial ASVs in gut tissue bacterial communities of spruce budworm larvae of different sexes. Only ASVs representing more than 0.01% of the entire analysis are shown in the graph. 61

Figure 2.6: Mean relative abundance of ASV1 and ASV5 for spruce budworm bacterial communities at sampling sites. 62

Figure 2.7: Boxplot showing the comparison between bacterial communities of spruce budworm whole larvae and gut tissue bacterial communities in terms of Shannon index. P-values of Mann-Whitney-Wilcoxon tests between larval and gut tissue community Shannon index are shown over each boxplot. 64

Figure 2.8: NMDS ordination of spruce budworm bacterial communities on Bray-Curtis dissimilarity for each site, whole larvae are shown as diamonds and gut tissue samples as circles. 68% confidence intervals ellipses representing sites are shown overlaid on the ordination. 65

Figure 2.9: Boxplot showing the comparison between spruce budworm bacterial communities of whole larvae and gut tissues in terms of distance to group centroid in non-Euclidean distance measured by betadisper (Anderson et al., 2006). Groups were defined as combinations of sample type and site factors. P-values of Mann-Whitney-Wilcoxon tests between larval and gut tissue community distances to centroid are shown over each boxplot. 66

Figure 2.10: DeSEQ2 discriminant analysis of differentially abundant ASVs in spruce budworm bacterial communities of gut tissues and whole larvae. ASVs found in both larval and gut tissue communities are identified as diamonds whereas ASVs found in only one sample type are identified as circles. 68

Figure 3.1: Co-occurrence network of bacterial communities in spruce budworm larvae and gut tissues collected from Quebec and Labrador and Newfoundland. Co-occurrence between every pair of genera was calculated with sparcc with default parameters by estimating Pearson correlations between log-transformed abundances. Community tables were randomized by permutation and p-values were calculated for each pair of genera with a setting of 1000 bootstraps to measure whether co-occurrence values observed differed from co-occurrence in randomly generated communities. Full edges indicate positive significant ($p < 0.05$) co-occurrences between genera. Dotted edges indicate negative significant ($p < 0.05$) co-occurrences between genera. Node sizes are proportional to relative abundance. Node colors represent phylum-level taxonomy. Genera whose co-occurrence is present in networks of both groups are labeled. 87

Figure 3.2: Proportion of positive and negative co-occurrences, network size, order and linkage density for networks of bacterial taxa collected from spruce budworm in Quebec. Co-occurrence networks were based on co-occurrence between every pair of genera calculated with sparcc by estimating Pearson correlations between log-transformed abundances. 89

Figure 3.3: Graph representation of each genus for each network in networks of spruce budworm larvae and gut tissues, where vertices are shaded depending on their ranks in the null model distribution and their size is proportional to relative abundance. Only genera outside the 95% confidence interval are labeled. 90

Figure 3.4: Co-occurrence network of bacterial communities in spruce budworm larvae and gut tissues feeding on two different host-tree species collected from Quebec and Labrador and Newfoundland. Co-occurrence between every pair of genera was calculated with sparcc with default parameters by estimating Pearson correlations between log-transformed abundances. Community tables were randomized by permutation and p-values were calculated for each pair of genera with a setting of 1000 bootstraps to measure whether co-occurrence values observed differed from co-occurrence in randomly generated communities. Full edges indicate positive significant

(p < 0.05) co-occurrences between genera. Dotted edges indicate negative significant (p < 0.05) co-occurrences between genera. Node sizes are proportional to relative abundance. Node colors represent phylum-level taxonomy. 92

Figure 3.5: Violin plot of measured network metrics compared to null model distribution in networks of spruce budworm larvae and gut tissues feeding on two different host-tree species. The white shapes are the null distributions, the colored lines are the observed values for our networks (purple for the balsam fir gut tissue network, cyan for the black spruce gut tissue network, orange for the balsam fir whole larvae network and green for the black spruce whole larvae network). Yellow shading areas represent values outside 95% confidence intervals (with 999 null models, ranks either under 25 or over 975 out of 1000). 93

Figure 3.6: Graph representation of each genus for each network in networks of spruce budworm larvae and gut tissues feeding on two different host-tree species, where vertices are shaded depending on their ranks in the null model distribution and their size is proportional to relative abundance. Only genera outside the 95% confidence interval are labeled. 94

Figure 3.7: Co-occurrence network of bacterial communities in spruce budworm gut tissues from larvae of different sexes collected from Quebec. Co-occurrence between every pair of genera was calculated with sparcc with default parameters by estimating Pearson correlations between log-transformed abundances. Community tables were randomized by permutation and p-values were calculated for each pair of genera with a setting of 1000 bootstraps to measure whether co-occurrence values observed differed from co-occurrence in randomly generated communities. Full edges indicate positive significant (p < 0.05) co-occurrences between genera. Dotted edges indicate negative significant (p < 0.05) co-occurrences between genera. Node sizes are proportional to relative abundance. Node colors represent phylum-level taxonomy. 96

Figure 3.8: Violin plot of measured network metrics compared to null model distribution in networks of spruce budworm gut tissues from larvae of different sexes. The white shapes are the null distributions, the colored lines are the observed values for our networks (light blue for the female gut tissue network and red for the male gut tissue network). Yellow shading areas represent values outside 95% confidence intervals (with 999 null models, ranks either under 25 or over 975 out of 1000). 97

Figure 3.9: Graph representation of each genus for each network in networks of spruce budworm gut tissues from larvae of different sexes, where nodes are shaded depending on their ranks in the null model distribution and their size is proportional to relative abundance. Only genera outside the 95% confidence interval are labeled..... 98

LISTE DES TABLEAUX

Table 1.1: Variation in spruce budworm bacterial communities as a function of sampling site and host tree species, tested using a Bray-Curtis dissimilarity-based PERMANOVA on ASV, genus, and family-level community tables. Host tree species was nested within site which was nested within region for this analysis. 36

Table 2.1: Variation in spruce budworm bacterial communities of gut tissue communities as a function of sampling site and larval sex tested using a Bray-Curtis dissimilarity-based PERMANOVA. 57

Table 2.2: Variation in the Shannon index of spruce budworm gut tissue bacterial communities as a function of larval sex and sampling site tested using an ANOVA on a generalized linear model fitted with the “gamma” distribution.. 60

Table 2.3: Variation in spruce budworm bacterial community composition of gut tissue and larval communities as a function of sample type, sampling site and host-tree species, tested using a Bray-Curtis dissimilarity-based PERMANOVA. 66

Table 2.4: Variation in spruce budworm bacterial community composition of gut tissue and larval samples as a function of sampling site and host-tree species nested within site, tested using a Bray-Curtis dissimilarity-based PERMANOVA, modeled separately for each sample type. 67

Table 3.1: Proportion of positive and negative co-occurrences, network size, order and linkage density for networks of bacterial taxa collected from spruce budworm in Quebec. Co-occurrence networks were based on co-occurrence between every pair of genera calculated with sparcc by estimating Pearson correlations between log-transformed abundances. 88

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

TBE	Tordeuse des Bourgeons de l'Épinette
ASVs	*Amplicon Sequence Variants*
ANOVA	*Analysis of Variance*
PERMANOVA	*Permutational Analysis of Variance*
PCoA	*Principal Coordinates Analysis*
MEMs	*Moran Eigenvector Maps*
NMDS	*Non-metric Multidimensional Scaling*

RÉSUMÉ

Il existe un lien important entre le microbiome et l'hôte dans une grande variété d'organismes, parmi lesquels sont les humains, des animaux et des plantes. Chez les insectes, le microbiome joue un rôle dans la défense contre les pathogènes, l'obtention de nutriments à partir de l'alimentation et la détermination sexuelle, entre autres. La composition et le rôle du microbiome varient de façon importante à travers la classe des insectes, et nos connaissances sur les facteurs qui déterminent la composition de ces communautés microbiennes sont insuffisantes. Chez les Lépidoptères, il a été théorisé que la larve ne peut entretenir une communauté résidente dans son intestin, et donc que la majorité du microbiome de l'insecte est transitoire et provient majoritairement de l'alimentation. On en sait très peu sur le microbiome de la tordeuse des bourgeons de l'épinette (TBE), *Choristoneura fumiferana*, un des Lépidoptères les plus destructeurs en Amérique du Nord. La larve de la TBE attaque principalement les populations de sapin et d'épinette. Périodiquement, les populations de cet insecte augmentent drastiquement, provoquant une épidémie et causant la défoliation d'importantes populations d'arbres hôtes. L'insecte fait partie d'un réseau trophique comportant divers animaux et insectes qui ont un impact sur la dynamique de ses populations, et dont les communautés bactériennes pourraient être une partie méconnue. Il est donc important de caractériser les facteurs déterminant la composition et les associations bactériennes du microbiome de la TBE à la fois pour mieux comprendre l'assemblage de microbiomes de Lépidoptères et comme une première étape pour améliorer nos outils de contrôle cet insecte important à la fois économiquement et écologiquement.

Cette thèse présente un projet visant à caractériser les facteurs spatiaux, environnementaux et biologiques déterminant la composition des communautés bactériennes de la TBE et les associations bactériennes à l'œuvre dans ces communautés. Notre premier objectif était de mesurer l'impact de facteurs environnementaux, de la composition des communautés du feuillage consommé par la TBE, de l'espèce d'arbre hôte et de la structure spatiale sur la composition des communautés bactériennes de la TBE. Pour ce faire, nous avons caractérisé les communautés microbiennes de larves de TBE à travers l'aire d'épidémie au Québec dans les provinces avoisinantes sur différentes espèces d'arbre hôte. Notre deuxième objectif était de mesurer à quel point les communautés des intestins de TBE variaient selon le sexe de l'insecte et comment les communautés des intestins différaient des communautés retrouvées dans les larves entières. Nous avons donc observé comment

ces communautés différaient en termes de diversité taxonomique, de composition générale et de variabilité entre individus. Notre troisième objectif était de mesurer comment la variation entre intestins et larves entières, entre espèces d'arbre hôte et entre sexes larvaires influençait les réseaux d'associations microbiens de la TBE. À cette fin, nous avons caractérisé les patrons de co-occurrence bactériens du microbiome de la TBE et avons déterminé la variation en termes de structure et de taxons moyeux entre les groupes mentionnés précédemment.

Le premier chapitre de cette thèse a démontré que les communautés du feuillage, la structure spatiale et l'espèce d'arbre hôte sont des facteurs importants dans la composition du microbiome de la TBE, mais qu'une importante variation spatiale demeure inexpliquée. Le deuxième chapitre a permis d'observer que malgré le fait que la composition du microbiome de l'intestin de la TBE ne varie que très peu selon le sexe larvaire, la diversité taxonomique et l'abondance de taxons spécifiques varient. Nous avons aussi démontré que les communautés de larves entières sont beaucoup plus diverses et variables entre individus que les communautés des intestins, mais que ces dernières sont beaucoup plus sensibles à l'effet du site d'échantillonnage. Lors du troisième chapitre, nous avons démontré que les réseaux de co-occurrence variaient en termes en structure, de densité ou de taxons moyeux selon l'espèce d'arbre-hôte, le sexe de l'insecte ou de si la communauté provient de larves entières ou des intestins de TBE. Le projet présenté dans cette thèse avance considérablement notre compréhension de l'assemblage des communautés bactériennes associées à la TBE et génère d'importantes hypothèses envers l'importance du microbiome pour l'insecte et l'amélioration des outils pour le contrôler.

Mots clés : Microbiome, Insectes défoliateurs, Écologie microbienne, Tordeuse des bourgeons de l'épinette, Réseaux de co-occurrence

ABSTRACT

There is an important link between the microbiome and the host in a wide variety of organisms, including humans, animals and plants. In insects, the microbiome plays a role in defending against pathogens, obtaining nutrients from food and determining sex, among others. The composition and role of the microbiome vary widely across insects, and our knowledge of the factors that determine the composition of these microbial communities is insufficient. In Lepidoptera, it has been theorized that the larva cannot support a resident community in its gut, and therefore that the majority of the insect's microbiome is transient and comes mainly from food. Very little is known about the microbiome of the eastern spruce budworm, *Choristoneura fumiferana*, one of the most destructive Lepidoptera in North America. The spruce budworm mainly attacks fir and spruce populations. Periodically, the populations of this insect increase drastically, causing an outbreak and causing defoliation of large populations of host trees. The insect is part of a food web comprising various animals and insects that have an impact on its population dynamics, and of which bacterial communities may be an unrecognized part. It is therefore important to characterize the factors determining the composition and bacterial associations of the spruce budworm microbiome both to better understand the assembly of Lepidopteran microbiomes and as a first step to improve our tools to control this insect both economically and environmentally very important.

This thesis presents a project aiming to characterize the spatial, environmental and biological factors determining the composition of the bacterial communities of the spruce budworm and the bacterial associations at work in these communities. Our first objective was to measure the impact of environmental factors, the composition of the foliage communities consumed by the spruce budworm, the host tree species and the spatial structure on the composition of the bacterial communities of the spruce budworm. To do this, we characterized the microbial communities of spruce budworms across the outbreak area in Quebec in neighboring provinces on different host tree species. Our second objective was to measure how much the communities of budworm gut tissues varied with the sex of the insect and how the communities of the gut tissues differed from the communities found in whole larvae. We therefore observed how these communities differed in terms of taxonomic diversity, general composition and variability between individuals. Our third objective was to measure how variation between gut tissues and whole larvae, between host tree species and between larval sexes influences the networks of microbial associations of spruce

budworm. To this end, we characterized the bacterial co-occurrence patterns of the spruce budworm microbiome and determined the variation in structure and hub taxa between the previously mentioned groups.

The first chapter of this thesis demonstrated that foliage communities, spatial structure and host tree species are important factors in the composition of the spruce budworm microbiome, but that significant spatial variation remains unexplained. The second chapter showed that despite the fact that the composition of the gut tissues microbiome of the spruce budworm varies little with larval sex, taxonomic diversity and abundance of specific taxa varies with sex. We have also shown that whole larvae communities are much more diverse and variable between individuals than intestinal communities, but that the latter are much more sensitive to the effect of the sampling site. In the third chapter, we demonstrated that bacterial co-occurrence networks vary in terms of structure, density or hub taxa depending on the host tree species, insect sex or whether the community is from whole larvae or budworm gut tissues. The project presented in this thesis considerably advances our understanding of the assembly of bacterial communities associated with the spruce budworm and generates important hypotheses regarding the importance of the microbiome for the insect and the improvement of currently available tools to control it.

Keywords : Microbiome, Defoliating pests, Microbial ecology, Spruce budworm, Co-occurrence networks

INTRODUCTION

Dans cette thèse je présente une étude des communautés bactériennes reliées à la tordeuse des bourgeons de l'épinette, *Choristoneura fumiferana*, dans un contexte de forêt naturelle touchée par une épidémie d'insectes défoliateurs. Le but de ce projet est d'établir l'importance relative de plusieurs facteurs biologiques et abiotiques sur la composition et les réseaux d'interactions du microbiome de *C. fumiferana*.

0.1 Définition du microbiome

Le terme microbiome a été inventé par Joshua Lederberg pour décrire la communauté de micro-organismes qui partagent un espace commun (Lederberg & McCray, 2001). Depuis le développement du séquençage à haut débit, la quantification de la diversité microbienne est devenue de moins en moins coûteuse, ce qui a mené à la caractérisation de la diversité taxonomique et fonctionnelle des communautés microbiennes dans une large gamme d'hôtes (Dubilier et al., 2015). Comme c'est souvent le cas dans les domaines de recherche en lien avec la santé humaine, l'essentiel de l'intérêt et des financements relatifs à ces nouvelles méthodes a été orienté vers les microbiomes humains. Ainsi, une partie importante de ces études porte sur les interactions homme-microbes et a généré une abondance de données sur les impacts des microbes résidant dans le corps humain (Bäckhed et al., 2012; NIH Human Microbiome Portfolio Analysis Team, 2019). Il a été démontré que les microbes dans l'intestin humain ont un impact sur la digestion, hydrolysant les glucides complexes de la matière végétale (Cantarel et al., 2012). Le microbiome humain contribue également à la maturation et à la régulation du système immunitaire,

ce qui peut être observé dans le cas des maladies inflammatoires de l'intestin, qui sont causées par la réponse excessive d'un système immunitaire non préparé à certaines bactéries (Kostic et al., 2014). Les microbes associés à l'homme sont aussi reconnus pour leur effet sur la régulation hormonale et la prévalence de nombreuses maladies humaines comme l'obésité, l'asthme et le cancer (Cho & Blaser, 2012; Mohajeri et al., 2018). Cette relation importante entre l'hôte et son microbiome s'applique également aux microbiomes trouvés dans les systèmes non humains. Plusieurs mammifères sont associés avec des communautés microbiennes contribuant à l'obtention de nutriments à partir de leur diète (Brulc et al., 2009; Zhu et al., 2011). Il a été démontré que les microbes associés aux arbres peuvent augmenter la productivité de leur hôte, et potentiellement de toute une forêt (Fürnkranz et al., 2008). En raison de l'importance des microbes pour ces écosystèmes, comprendre comment leurs communautés s'assemblent permettrait de mieux comprendre comment ces écosystèmes fonctionnent dans leur ensemble (Six, 2013). La majorité des microbiomes étudiés sont ceux de mammifères, et il en reste encore beaucoup à découvrir sur les microbiomes d'animaux importants économiquement et écologiquement comme les insectes.

0.2 Les microbiomes d'insectes

Les insectes sont l'une des classes les plus diverses d'organismes vivants (Basset et al., 2012). En conséquent, il n'est pas surprenant que les microbiomes des insectes varient considérablement en composition et en fonction parmi les différentes espèces d'insectes (Dillon & Dillon, 2004; Moran et al., 2019). Cet variation peut être expliquée par le fait que les insectes ont des habitudes alimentaires très diverses et que la majeure partie de leur microbiome réside dans l'intestin (Engel & Moran, 2013). Les détritivores comme les termites, les grillons et les coléoptères possèdent les microbiomes les plus abondants en termes de cellules bactériennes (Cazemier et al., 1997), ce qui indiquerait que les traits et les rôles écologiques des insectes

peuvent être déterminants du type de communauté microbienne qui les colonisent. Le microbiome peut apporter de multiples avantages à son hôte et cela a été bien étudié dans le monde des insectes (Engel & Moran, 2013; Moran et al., 2019). L'un des aspects les mieux étudiés du microbiome des insectes est son effet sur la digestion. En effet, les termites sont bien connus pour être dépendant de protozoaires, champignons et bactéries pour digérer la matière ligneuse (Breznak, 1982; Engel & Moran, 2013; Huang et al., 2015; Peterson & Scharf, 2016; Rossmassler et al., 2015; Schäfer et al., 1996; Scully et al., 2013). Les bactéries des termites participent à la dégradation de la cellulose et de l'hémicellulose ainsi qu'à la fixation de l'azote (Warnecke et al., 2007). De nombreuses autres espèces d'insectes n'ont pas besoin de bactéries pour digérer leur nourriture, mais les bactéries qui leur sont associés leur procurent d'autres avantages (Calderón-Cortés et al., 2011). Ces avantages pourraient aussi parfois être hérités par la progéniture de l'insecte et favoriser sa valeur sélective (A. H. Smith et al., 2015). Par exemple, chez le puceron du pois, la symbiose microbiome-insecte influence les interactions avec ses ennemis naturels, menant potentiellement à une réduction des probabilités d'attaque réussie par les prédateurs parasitoïdes (A. H. Smith et al., 2015). De plus, il a été démontré que le microbiome aide à protéger les abeilles, les moustiques et les grillons contre d'autres microbes pathogènes et parasites (Engel & Moran, 2013). Chez la drosophile, le microbiome contribue au développement et au maintien de l'épithélium intestinal (Buchon et al., 2013). Finalement, le microbiome pourrait jouer un rôle dans la détoxification d'insecticides et la production de phéromones chez d'autres insectes (Dillon & Dillon, 2004; Kikuchi et al., 2012).

0.3 La tordeuse des bourgeons de l'épinette et son impact

La tordeuse des bourgeons de l'épinette (TBE), *Choristoneura fumiferana*, est l'un des ravageurs forestiers les plus destructeurs en Amérique du Nord (Martineau, 1985). La TBE est un papillon indigène de la famille des Tortricidae dont les larves se

nourrissent de bourgeons et de jeunes aiguilles de conifères. Le cycle de vie de *C. fumiferana* dure un an et commence à la fin de l'été, lorsque les papillons pondent des masses d'œufs sur les aiguilles. Les œufs éclosent quelques semaines plus tard et les larves s'abritent pour passer l'hiver en diapause en tant que larves de deuxième stade. Au printemps, les larves émergent et se déplacent vers les jeunes bourgeons où elles tissent un abri en soie pour les protéger pendant qu'elles se nourrissent pendant environ six semaines jusqu'à ce qu'elles se transforment en six stades larvaires et finalement en pupes (Martineau, 1985). Son arbre hôte de prédilection est le sapin baumier (*Abies balsamea*), pour lequel l'émergence des bourgeons coïncide avec l'émergence des larves de leur diapause. L'épinette noire (*Picea mariana*), rouge (*Picea rubens*) et blanche (*Picea glauca*) sont également des hôtes importants de la TBE, même si leurs bourgeons émergent plus tard (Hennigar et al., 2008).

Les épidémies de TBE apportent des changements complexes au fonctionnement de l'écosystème forestier (Burleigh, 1998). La destruction de feuillage par la TBE réduit le nombre de graines potentielles pour l'hôte, ce qui réduit la productivité et modifie la distribution des essences de conifères dans la forêt boréale. Les arbres survivants passent par une phase de stagnation, prenant encore de la place et des ressources mais avec une productivité réduite (Burleigh, 1998). À faible densité larvaire, les bourgeons frais sont mangés et l'arbre ne perdra qu'une partie du feuillage de l'année en cours. Lorsque la densité des populations de TBE est élevée, les larves se nourrissent également d'aiguilles plus âgées, l'affaiblissant considérablement et pouvant entraîner sa mort après 4 à 6 ans de défoliation (MacLean & Piene, 1995). Ces décès sont causés par le désavantage compétitif des arbres défoliés pour l'apport de lumière, d'eau et de nutriments (Pothier & Mailly, 2006). Additionnellement, les populations d'arbres dévastées peuvent potentiellement servir de combustible pour les incendies de forêt, augmentant éventuellement leur gravité et leur occurrence dans les années suivantes (James et al., 2017; Stocks, 1987). Des épidémies se produisent environ tous les 30 ans lorsque les populations atteignent des densités extrêmement

élevées (Jardon et al., 2003). Une épidémie est actuellement en cours au Québec, ayant déjà défolié plus de 13 500 000 hectares de forêt, avec l'épicentre dans la région de la Côte Nord du Québec (Ministère des Forêts, 2020). Plusieurs années consécutives de défoliation intense entraînent des pertes économiques massives dans la production de bois (MacLean et al., 2002). Par exemple, sans programmes de contrôle, il est projeté sont que la TBE sera responsable de pertes d'environ quatre milliards de dollars pour l'industrie forestière au Nouveau-Brunswick au cours des 35 prochaines années (Chang et al., 2012). La principale méthode de contrôle biologique utilisée au Canada est l'épandage de BT, une préparation à base de la bactérie *Bacillus thurigiensis*, qui lorsqu'elle sporule produit une toxine qui perfore l'intestin des Lépidoptères et mène à leur mort (Palma et al., 2014).

0.4 Les microbiomes de Lépidoptères

Le microbiome de certains Lépidoptères a déjà été caractérisé, principalement pour des insectes économiquement importants comme le ver à soie ou des ravageurs défoliateurs (Anand et al., 2010). Des études sur les larves de papillons ont montré que la majorité des bactéries qui composent la communauté microbienne de leur intestin proviennent de leur nourriture (Broderick et al., 2004; Hammer et al., 2017). Les insectes lépidoptères sont également connus pour avoir un microbiome de faible diversité par rapport aux autres insectes (Engel & Moran, 2013). Leur régime alimentaire se compose principalement de matières végétales, qui contiennent des molécules qui acidifient leur tube digestif et rendent plus difficile la colonisation par les bactéries (Berenbaum, 1980; Harrison, 2001). Le tube digestif des Lépidoptères est également dépourvu de toute adaptation connue destinée à protéger ou à abriter des bactéries, que l'on peut trouver chez d'autres espèces d'insectes. Ces propriétés suggèrent qu'il est plus difficile pour les Lépidoptères de maintenir un microbiome stable et diversifié. Nous supposons que parce que seules les larves ont besoin de l'intestin (Capinera, 2008), c'est-à-dire que les adultes ne se nourrissent pas, la

pression sélective a abouti à une structure intestinale aussi simple que possible. Une autre possibilité est que les microbes nécessaires à la digestion des aliments sont omniprésents sur le feuillage dans la nature et qu'aucune adaptation n'est nécessaire pour retenir ces microbes. Ces théories supportent l'idée que les intestins de larves de lépidoptères ne peuvent héberger une communauté microbienne résidente, comme il a déjà été démontré en laboratoire (Hammer et al., 2017). Par contre, cette hypothèse n'a toujours pas été testée empiriquement en milieu naturel.

Les connaissances sur le sujet du microbiome de la TBE sont très limitées. Des expériences de culture ont démontré la présence d'*Enterococcus mundtii*, *Staphyloccocus succinus*, *Bacillus subtilis*, *Pseudomonas* et *Paenibacillus* et que l'antibiotique présent dans la diète artificielle typique (auréomycine) n'est pas assez puissant et n'affecte pas un spectre assez large pour éliminer complètement le microbiome intestinal de la larve (Frankenhuyzen et al., 2010). Il est encore débattu si le mécanisme d'action du BT implique le microbiome de l'intestin de la tordeuse pour la tuer par infection de l'hémocèle ou si la perforation de l'intestin suffit à tuer l'insecte (Broderick et al., 2006, 2009; Frankenhuyzen et al., 2010). Des études de séquençage 16S ont montré que le microbiome de l'intestin de la larve était dominé par le genre *Pseudomonas* et que l'environnement d'élevage et/ou le régime alimentaire affecte la composition de ces communautés microbiennes (Landry et al., 2015). Cette étude a également démontré que les larves élevées avec une diète artificielle contenant de l'auréomycine possédaient un microbiome plus diversifié que les larves échantillonnées sur le terrain. Landry et al. (2015) ont aussi observé qu'en termes de composition du microbiome, les échantillons de chaque groupe expérimental se divisaient en deux groupes de tailles similaires, qu'ils suggèrent sont relatifs au sexe des larves. Une expérience d'élevage en laboratoire a démontré que le microbiome de l'intestin de la TBE n'avait pas un impact détectable sur le poids ni la survie de son hôte (Schwarz et al., 2018).

0.5 Pourquoi étudier le microbiome de la TBE

Malgré ces informations sur le microbiome associé à la TBE, nous en savons encore très peu sur les processus qui conduisent à l'assemblage de ces communautés et l'importance du microbiome dans un contexte d'épidémie sur une grande échelle spatiale. Ce projet vise à mieux comprendre quels processus écologiques déterminent les patrons de communautés microbiennes observés dans la nature et à mieux comprendre la relation entre la TBE, son microbiome, l'intensité de défoliation et la dynamique des épidémies. Cela pourrait également conduire à terme à l'élaboration de nouveaux moyens de lutte contre les parasites ou à l'amélioration de ceux qui existent déjà. En effet, comme mentionné précédemment, le principal insecticide (BT) utilisé pour lutter contre la TBE pourrait avoir son efficacité modulée par le microbiome intestinal de la tordeuse des bourgeons (Frankenhuyzen et al., 2010) et une meilleure compréhension de la communauté microbienne associée à la TBE pourrait conduire à un formulation BT plus efficace. Enfin, la compréhension des facteurs déterminant la composition du microbiome de la TBE et des bactéries qui sont associées à la défoliation peut potentiellement aider à identifier les populations d'arbres vulnérables ou résistantes. Les agences de gestion de protection des forêts pourraient ensuite utiliser ces informations pour décider quelles populations d'arbres à protéger contre la TBE. Finalement, le système de la TBE présente une opportunité importante pour mieux comprendre comment les communautés associées aux Lépidoptères et leur alimentation sont distribuées spatialement à grand échelle.

0.6 Processus dirigeant la formation de communautés microbiennes

Il existe une énorme diversité de microbiomes et une problématique importante en écologie microbienne est de déterminer quels processus déterminent cette diversité. Vellend (2010) suggère que la structure des communautés est déterminée

principalement par quatre processus : la sélection, la dérive, la dispersion et la spéciation. Cette théorie provient principalement de l'écologie végétale mais demeure pertinente en ce qui a trait aux communautés microbiennes (Nemergut et al., 2013). La sélection, qui est un processus qui favorise les taxons les mieux adaptés à leur environnement, est constatable dans les importantes différences de communautés bactériennes d'un type d'habitat à un autre (J. B. H. Martiny et al., 2006; Nemergut et al., 2013). Le filtrage environnemental, qui est le processus selon lequel des propriétés d'un environnement préviennent des taxons spécifiques de le coloniser (Kraft et al., 2015), joue un rôle important dans la sélection. La dérive est décrite comme des variations aléatoires des niveaux de population d'une communauté et a un effet plus important dans les petites communautés non diversifiées où la sélection n'est pas un facteur important (Nemergut et al., 2013). Les microbiomes relativement simples de certains organismes vivants pourraient en être un bon exemple, comme dans les intestins des iguanes où la dérive écologique a un impact important sur la communauté (Lankau et al., 2012). La dispersion est définie comme le mouvement des espèces d'un habitat à un autre, ce qui est difficile à observer chez les microbes en raison de leur taux de reproduction élevé, de leur taille microscopique et de leurs communautés denses (Nemergut et al., 2013). Parce que le concept d'espèce bactérienne est encore mal défini (Brenner et al., 2005) et avec l'occurrence de transfert génétique horizontal (Willey, 2014), le processus de spéciation dans le cas de communautés microbiennes pourrait être mieux défini comme diversification, représentant l'apparition de nouveaux génotypes (Nemergut et al., 2013). Cette apparition de nouveaux génotypes peut être stimulée par la tendance de certains organismes aux fonctions similaires cohabitant à « partager » des ressources en se spécialisant dans des voies métaboliques légèrement similaires mais différentes, appelée partitionnement de niche (Griffin & Silliman, 2012).

0.7 Les facteurs de variation spatiale et les « Moran eigenvectors maps »

On en sait très peu sur comment les microbiomes d'animaux varient à travers l'espace car très peu d'études s'intéressent au lien entre le microbiome et les facteurs géographiques. La configuration spatiale des populations hôtes du microbiome pourrait affecter la composition du microbiome en affectant la dispersion à partir d'un réservoir d'espèces régionales (Nemergut et al., 2013; Vellend, 2016). Chez les végétaux, on sait que l'emplacement géographique est un des facteurs déterminant la composition des communautés présentes sur le feuillage (Laforest-Lapointe et al., 2016). Comme mentionné précédemment, ce que nous savons présentement des microbiomes de Lépidoptères suggère que la diète de l'insecte est la principale source de dispersion bactérienne (Hammer et al., 2017), donc la variation spatiale sur le feuillage consommé par la TBE se traduira forcément en une variation spatiale des communautés microbiennes de la TBE elle-même. C'est pour combler ce manque de connaissances qu'il est important de déterminer comment le microbiome de la TBE varie spatialement sur une grande échelle. Le test de Mantel a longtemps été la principale méthode pour tester l'impact de distances géographiques sur des communautés, mais nous proposons une méthode qui a été démontrée comme étant beaucoup plus appropriée pour répondre à nos questions, les « Moran eigenvectors maps » (MEMs) (Harmon & Glor, 2010; Legendre et al., 2005, 2015; Legendre & Fortin, 2010; Lisboa et al., 2014). Plus précisément, cette méthode nous permet de quantifier la structure spatiale de sites d'échantillonnage à l'aide de vecteurs orthogonaux qui quantifient l'autocorrélation spatiale dans les données géospatiales, appelés « Moran eigenvectors » (Borcard et al., 2011). L'analyse MEM permet de détecter des structures spatiales qui ne sont pas ou peu observables dans une matrice de distances géographiques et de les utiliser comme variables prédictives dans la modélisation multivariée souvent utilisée pour l'analyse de communautés (Legendre et al., 2015).

0.8 Les facteurs environnementaux

Certains facteurs abiotiques tels que l'humidité, la température et les précipitations ont un effet sur la croissance bactérienne (Balser et al., 2010). Plus précisément, ils pourraient agir comme des filtres qui permettent uniquement à un sous-ensemble du pool d'espèces régionales de se disperser et de s'établir dans un habitat (Kraft et al., 2015). Il a déjà été démontré comme que le climat était un facteur déterminant de la structure des communautés écologiques (Hernández-Gómez, 2020; Sepulveda & Moeller, 2020). Dans le contexte de la TBE, les différences climatiques et météorologiques à travers la superficie touchée par la TBE pourraient donc moduler la sélection et l'établissement de populations microbiennes qui vont s'associer à l'insecte.

0.9 Les différents arbres-hôtes

Outre les facteurs spatiaux et environnementaux, l'espèce d'arbre hôte pourrait être un des facteurs importants de variation dans le microbiome de la TBE. L'espèce pourrait avoir un impact pour deux raisons différentes: (1) parce que les divers feuillages transfèrent différentes communautés à la TBE lors de leur contact et de leur ingestion (Laforest-Lapointe et al., 2016); (2) parce que la nature du feuillage lui-même affecte les communautés associées à la TBE (Pang, 2010). Les différentes espèces d'arbres ont des feuilles très différentes en termes de nutriments, de surface foliaire spécifique et de résistance mécanique, en plus de contenir différents produits chimiques affectant la croissance bactérienne (Pang, 2010). La variation de maturité des aiguilles causée par la différence de phénologie entre les arbre-hôtes pourrait aussi jouer un rôle dans la sélection de microbes à être introduits dans la TBE et la facilité avec lesquels ils peuvent se propager (Hennigar et al., 2008). Chez d'autres espèces d'arbres retrouvés au Québec, l'espèce de l'arbre est effectivement un des

majeurs facteurs déterminant le type de microbiome se retrouvant sur ses feuilles, ce qui est probablement le cas chez les hôtes de la TBE, qui offriraient différents microbiomes à l'insecte (Laforest-Lapointe et al., 2016).

0.10 Différences du microbiome reliées au sexe de l'hôte

Depuis la popularisation de la caractérisation de microbiomes, le sexe de l'hôte a été considéré comme un facteur potentiel de variation du microbiome (Jašarević et al., 2016). Il a été démontré que le microbiome humain est modulé par les hormones sexuelles et d'autres différences physiologiques entre les sexes affecte la composition et la fonction du microbiome dans l'intestin et a des impacts importants sur l'immunité et la santé globale (Jašarević et al., 2016; Valeri & Endres, 2021; Vemuri et al., 2019). Des différences de comportement peuvent aussi être la cause de variation de la composition du microbiome, comme chez les souris (Bridgewater et al., 2017). Chez les insectes, la bactérie *Wolbachia* est reconnue pour jouer un rôle dans la différenciation sexuelle chez plusieurs espèces d'insectes, jouant même un rôle dans le taux de mortalité des mâles (Graham & Wilson, 2012; Pattabhiramaiah et al., 2012; J. A. Russell et al., 2012). Les communautés microbiennes habitant le système reproducteur de moustiques et le système digestif de drosophiles sont enrichis en différents taxons microbiens selon le sexe de l'insecte (Han et al., 2017; Segata et al., 2016). Chez d'autres insectes, des différences sexuelles de taille, de comportement de ponte, de résistance aux parasites et de régime alimentaire entre les sexes pourraient mener à des différences de composition du microbiome comme observé chez les mammifères (Baker et al., 1987; Fanson et al., 2013; Nakashima & Hirose, 2003; Willink & Svensson, 2017).

0.11 Variation du microbiome à travers l'hôte

L'intestin des insectes héberge des communautés microbiennes très diverses et abondantes formant la majeure partie leur microbiome (Engel & Moran, 2013; Moran et al., 2019), mais des communautés bactériennes importantes sur le plan écologique et fonctionnel ont été retrouvés dans d'autres parties de l'insecte (Otti, 2015). En fait, bien que nous sachions que différents tissus et parties du corps abritent des communautés microbiennes distinctes chez les mammifères (Dekaboruah et al., 2020; Stehlikova et al., 2019), on ne sait généralement pas comment la structure de la communauté microbienne varie entre les parties du corps pour les insectes, en particulier pour les Lépidoptères. Les communautés sur la surface d'un insecte pourraient potentiellement être très sensibles aux conditions environnementales, car elles sont susceptibles d'être colonisées par des microbes se dispersant à partir de l'environnement de l'insecte. D'autre part, les communautés intestinales devraient varier principalement avec le régime alimentaire de l'insecte, qui est la principale source de colonisation bactérienne mais aussi l'un des principaux facteurs déterminant les conditions abiotiques présentes dans l'intestin de l'insecte (Hammer et al., 2017; Moran et al., 2019). Comme on peut l'observer chez le termite, qui héberge des communautés spécialisées dans différents compartiments de son intestin, différents microbes pourraient avoir co-évolué avec l'organe qui leur est associé pour contribuer aux fonctions de cette partie du corps (Breznak, 1982; Engel & Moran, 2013; Huang et al., 2015; Peterson & Scharf, 2016; Rossmassler et al., 2015; Schäfer et al., 1996; Scully et al., 2013). De ce que l'on en sait, chez les Lépidoptères, le microbiome des excréments et larves entières étaient de compositions extrêmement similaires (Hammer et al., 2014), mais significativement différentes du microbiome de segments d'intestin disséqués (Hammer et al., 2017), suggérant que les mesures basées sur les excréments et les larves entières des communautés microbiennes ne représentent pas forcément avec justesse les communautés intestinales de l'insecte.

0.12 Les réseaux de co-occurrence microbiens

Les réseaux sont des outils utiles pour visualiser les interactions et la co-occurrence au sein d'une communauté et permettent d'observer la complexité des communautés microbiennes (Dunne et al., 2002; Kishore et al., 2020; Layeghifard et al., 2017). Il est connu que les microbes qui font partie de communauté forment des relations mutuellement bénéfiques qui ne sont pas détectables en mesurant seulement leur abondance relative (Dalton et al., 2011; Murray et al., 2014). L'analyse de réseaux peut aussi donner un aperçu des mécanismes à l'origine des schémas d'abondance relative, qu'ils impliquent des facteurs abiotiques, les microbes eux-mêmes ou l'hôte (Agler et al., 2016). Les réseaux d'associations microbiennes sont généralement déduits en mesurant leur co-occurrence (van der Heijden & Hartmann, 2016). En d'autres termes, la corrélation entre les abondances relatives de chaque paire de taxons est testée et lorsque l'abondance de deux taxons est corrélée, on peut interpréter qu'ils interagissent potentiellement ensemble (Friedman & Alm, 2012). Ces interactions pourraient être positives (coopération ou symbiose) si deux taxons sont généralement retrouvés ensemble, ou bien négatives (compétition ou prédation) lorsque l'abondance d'un taxon réduit l'abondance d'un autre (Delmas et al., 2019). Par contre, malgré le fait que de nombreuses études supposent que les réseaux de co-occurrence microbienne sont la preuve d'interactions microbiennes réelles, il a été proposé que des problèmes méthodologiques et théoriques mettent en doute la pertinence de telles assumptions (Blanchet et al., 2020). La détection de réelles interactions écologiques dépendrait de l'environnement, de la présence d'autres interactions directes et indirectes dans une communauté (Blanchet et al., 2020) et de caractéristiques spécifiques du plan d'échantillonnage. Cependant, connaître les réseaux de co-occurrences microbiens demeure un outil utile pour caractériser les patrons de variation d'abondance et générer des hypothèses écologiques allant au-delà des tables d'abondances relatives. De plus, ce type d'analyse permet d'identifier

des taxons particulièrement centraux (taxons moyeux) qui co-occurrent avec un nombre important d'autres taxons, signifiant qu'ils ont potentiellement un effet important sur la communauté et leur hôte (Agler et al., 2016).

L'analyse des réseaux de communautés microbiennes est encore récente et les connaissances que nous possédons sur les réseaux microbiens dans la nature sont peu abondantes (Delmas et al., 2019; Layeghifard et al., 2017). Les réseaux bactériens dans les filtres biologiques à sable possédaient des taxons centraux très peu abondants, démontrant que le degré de co-occurrence n'était pas forcément corrélé avec l'abondance (Zhao et al., 2021). Dans le réseau microbien associé à *Arabidopsis thaliana*, l'hypothèse a été émise que c'est par l'entremise de ces taxons moyeux que l'hôte peut avoir un impact sur son microbiome (Agler et al., 2016). Les réseaux bactériens du sol se sont avérés beaucoup plus riches en cooccurrences positives lorsqu'ils subissaient du stress environnemental, et offraient une meilleure résistance à un pathogène lorsqu'ils étaient densément connectés (Hernandez et al., 2021; Legrand et al., 2019). De plus, ces réseaux variaient de façon importante en fonction de gradients spatiaux et environnementaux (Tu et al., 2020). Des patrons de co-occurrence complexes entre bactéries et champignons ont été retrouvées chez le foreur du pommier, *Agrilus mali*, parmi lesquels se trouve un forte corrélation positive entre les taxons *Kocuria* et *Corynebacterium* (Zhang et al., 2018).

0.13 Présentation de la thèse

L'objectif de ce projet de doctorat est de caractériser différents facteurs expliquant les variations du microbiome de la TBE en milieu naturel. Cette thèse est divisée en trois chapitres qui s'attardent sur: (1) les facteur spatiaux, environnementaux et biologiques qui déterminent la biogéographie du microbiome de la TBE; (2) comment les communautés bactériennes de l'intestin de la TBE varient selon son sexe et leur variation par rapport aux communautés de l'insecte en entier; (3) les

réseaux de co-occurrence présents dans le microbiome de la TBE et leur variation en fonction de l'arbre hôte, le sexe de l'insecte ou du type de communauté. Pour ce faire, des ensembles de séquences du gène codant pour la sous-unité ribosomale bactérienne 16S seront analysés en utilisant les plus récentes méthodes d'écologie microbienne, de bioinformatique et de science des réseaux.

0.13.1 Chapitre 1

Ce chapitre consiste à déterminer les facteurs déterminant l'assemblage et la biogéographie des microbiomes associés à la TBE. De manière générale, les communautés bactériennes de la tordeuse pourraient soit (a) provenir d'une combinaison de bactéries présentes dans le régime alimentaire de l'insecte, soit (b) provenir de l'environnement de l'insecte. Dans les deux cas, le réservoir de microbes disponibles dans l'environnement qui colonisent le microbiome de la TBE serait probablement déterminé par une variété de facteurs, notamment les conditions environnementales, l'autocorrélation spatiale, les microbes résidant sur le feuillage consommé par les larves et l'espèce de l'arbre hôte.

Objectifs

(1.1) Quantifier l'importance de quatre facteurs pour expliquer la variation des communautés bactériennes associées à la TBE et à quel point elles interagissent les unes avec les autres : 1) les conditions environnementales, 2) l'autocorrélation spatiale, 3) les communautés microbiennes du feuillage et 4) les espèces d'arbres hôtes ;

Hypothèses

H1.1 Il a été théorisé que les Lépidoptères ne peuvent héberger de communauté bactérienne résidente et donc qu'une majorité de cette communauté provient directement de l'alimentation (Hammer et al., 2017), mais ceci reste à être démontré chez la TBE. C'est pourquoi nous prédisons que la composition des communautés microbiennes dans le régime foliaire de la TBE est le principal déterminant de la composition des communautés bactériennes de la TBE.

H1.2 Puisque les communautés microbiennes du feuillage varient dans leur composition selon les espèces d'arbres hôtes et l'emplacement géographique (Laforest-Lapointe et al., 2016), l'autocorrélation spatiale et les espèces d'arbres hôtes sont des déterminants majeurs de la composition des communautés bactériennes du feuillage et de la TBE.

0.13.2 Chapitre 2

Ce chapitre consiste en l'exploration des différences de composition et de diversité du microbiome en fonction du sexe de la TBE et de si les communautés proviennent de l'intestin ou de la larve en entier. Répondre à ces questions permettrait d'en connaître plus sur le lien entre la physiologie de l'insecte et son microbiome. Plus précisément, ce projet va permettre de caractériser les différences écologiques entre l'intestin et les reste du corps de l'insecte en tant qu'habitats pour des communautés bactériennes et d'en savoir plus sur la nature des échantillons nécessaires pour étudier les microbiomes de Lépidoptères.

Objectifs

(2.1) Examiner comment les communautés bactériennes intestinales de la TBE varient en fonction du sexe de la larve.

(2.2) Caractériser les différences dans la composition et la diversité de la communauté bactérienne entre les intestins des larves et les larves entières.

Hypothèses

H2.1 Des différences en termes de diversité et composition dans le microbiome intestinal de la TBE existent entre les larves de différents sexes.

H2.2 Il existe une plus grande diversité dans les communautés larvaires entières que dans les communautés intestinales, car l'épiderme larvaire est exposé à l'environnement et à des dispersions aériennes potentielles alors que l'intestin est périodiquement inoculé par le même réservoir provenant de l'alimentation et sert de filtre environnemental.

H2.3 Les bactéries de l'intestin sont plus sensibles à l'espèce d'arbre-hôte car le régime alimentaire est potentiellement la principale source d'inoculation bactérienne dans l'intestin alors que les communautés des larves entières varient davantage selon les sites d'échantillonnage car elles sont affectées par l'environnement de l'insecte.

0.13.3 Chapitre 3

Ce chapitre vise à explorer les réseaux de co-occurrence microbiens impliqués dans le microbiome de la TBE et comment ils varient en fonction des facteurs étudiés dans les deux chapitres précédents. La caractérisation des réseaux de co-occurrence microbiens hébergés au sein des insectes pourrait aider à expliquer le niveau important de variabilité de leurs communautés microbiennes (Engel et al., 2012; Engel & Moran, 2013) ou encore conduire à une compréhension plus générale des facteurs régissant la composition du microbiome des insectes.

Objectifs

(3.1) Caractériser les différences de structure et de taxons moyeux entre des réseaux de cooccurrence de communautés d'intestins de larves et de larves entières.

(3.2) Caractériser les différences de structure et de taxons moyeux entre des réseaux de cooccurrence en fonction de l'espèce d'arbre hôte sur laquelle les larves se sont nourries.

(3.3) Caractériser les différences de structure et de taxons moyeux entre des réseaux de cooccurrence en fonction du sexe de la larve.

Hypothèses

H3.1 Étant donné que l'intestin est physiquement niché dans le corps larvaire, les co-occurrences microbiennes au sein des communautés d'intestins sont nichées dans celles des réseaux de larves entières.

H3.2 Les réseaux microbiens de larves entières impliquent des taxons plus diversifiés, ce qui entraînera un ordre de réseau plus élevé et un diamètre de réseau plus faible, car les larves entières sont beaucoup plus exposées à l'environnement que les intestins.

H3.3 Étant donné que deux taxons (*Rhodococcus* et *Mesorhizobium*) dominaient les communautés intestinales en termes d'abondance relative (avec 86 % des séquences attribuées à ces genres bactériens), ce sont des taxons moyeux dans les réseaux de co-occurrence intestinaux (Landry et al. Chapitre 2).

H3.4 Les réseaux de co-occurrence des larves entières varient plus en fonction de l'espèce d'arbre hôte que les réseaux de larves entières, en raison d'une plus grande sélectivité des intestins de la TBE.

H3.5 Les réseaux de communautés de larves de sexe différents présentent une structure et des taxons moyeux similaires, étant donné l'importante similarité de la composition générale des communautés de larves de différents sexes.

CHAPITRE I

SPRUCE BUDWORM BACTERIAL COMMUNITIES VARY AMONG SITES AND HOST TREE SPECIES IN A BOREAL LANDSCAPE

Cet article a été publié par le journal *Journal of Biogeography*.

Mathieu Landry, premier auteur.
Département des sciences biologiques, Université du Québec à Montréal,
141 Av. du Président-Kennedy, Montréal, Québec, H2X 1Y4
landry.mathieu.5@courrier.uqam.ca

Patrick M.A. James
Institute of Forestry and Conservation, University of Toronto,
33 Willcocks St, Toronto, Ontario, M5S 3B2
patrick.james@utoronto.ca

Daniel Kneeshaw
Département des sciences biologiques, Université du Québec à Montréal,
141 Av. du Président-Kennedy, Montréal, Québec, H2X 1Y4
kneeshaw.daniel@uqam.ca

Steven W. Kembel
Département des sciences biologiques, Université du Québec à Montréal,
141 Av. du Président-Kennedy, Montréal, Québec, H2X 1Y4
kembel.steven_w@uqam.ca

1.1 Résumé

Une question clé dans la biogéographie des microbiomes associés à l'hôte est de comprendre comment les facteurs spatiaux, environnementaux et associés à l'hôte influencent la composition de la communauté microbienne. Ici, nous avons cherché à évaluer l'importance relative de ces multiples facteurs sur la composition des communautés bactériennes associées à la tordeuse des bourgeons de l'épinette (TBE), *Choristoneura fumiferana*. Cette étude est une première étape vers la compréhension de la variation spatiale des communautés bactériennes d'un insecte dont la relation avec son microbiome n'est pas encore caractérisée. L'étude a été réalisée dans les forêts boréales du Québec, de Terre-Neuve et du Labrador où il y a une épidémie de TBE. Nous avons caractérisé les communautés bactériennes associées aux larves de la TBE dans l'est du Canada en utilisant le séquençage d'amplicon du gène bactérien de l'ARNr 16S. Nous avons ensuite quantifié dans quelle mesure la variation de ces communautés bactériennes pouvait s'expliquer par : (1) les conditions environnementales, (2) le microbiome du feuillage que les larves mangeaient, (3) les espèces d'arbres hôtes et (4) la structure spatiale telle que quantifiée à l'aide de *Moran Eigenvector Maps* (MEMs). Nous avons constaté que le microbiome de la TBE varie considérablement entre les sites et entre les espèces d'arbres hôtes. La composition des communautés bactériennes des larves de TBE était fortement corrélée avec la composition des communautés bactériennes prélevées sur des échantillons de feuillage appariés. La structure spatiale, les communautés bactériennes du feuillage et les espèces d'arbres hôtes expliquaient collectivement près d'un sixième de la variation des communautés bactériennes de la TBE. Les communautés bactériennes du feuillage, les espèces d'arbres hôtes et la structure spatiale semblent tous jouer un rôle différent dans la composition de la communauté bactérienne de la TBE, avec les communautés foliaires étant le facteur le plus important. Cette étude fait partie des premières études empiriques à démontrer la grande similitude entre la communauté bactérienne d'un insecte lépidoptère et celle de son régime alimentaire, soutenant l'hypothèse qu'il leur manque un microbiome résident.

1.2 Abstract

A key question in the biogeography of host-associated microbiomes is understanding how spatial, environmental, and host-associated factors influence microbial community composition. Here, we aimed to assess the relative importance of these multiple factors on the composition of bacterial communities associated with the eastern spruce budworm (*Choristoneura fumiferana*). This study is a first step towards understanding spatial variation in the bacterial communities of an outbreaking insect whose relationship with its microbiome is still uncharacterized. The study was carried out in boreal forests of Quebec, Newfoundland, and Labrador where there is an ongoing spruce budworm outbreak. We characterized bacterial communities associated with spruce budworm larvae in eastern Canada using bacterial 16S rRNA gene amplicon sequencing. We then quantified how much of the variation in these bacterial communities could be explained by: (1) environmental conditions, (2) the microbiome of foliage the larvae were eating, (3) host tree species and (4) spatial structure as quantified using Moran's Eigenvector Maps (MEMs). Budworm larval microbiomes varied significantly among sites and between host tree species. Larvae bacterial community structure was strongly correlated with the structure of bacterial communities taken from paired foliage samples. Spatial structure, foliage bacterial communities, and host tree species collectively explained almost one-sixth of the variation in budworm bacterial communities while environmental conditions did not explain variation on their own. Lepidopteran microbiomes primarily originate from the foliage diet. However, subtle differences in microbial communities between larvae and foliage suggest that some bacteria establish and grow in the budworm microbiome, and that dispersal of bacteria from sources other than foliage as well as differences in environmental filtering between larval bodies and foliage play a role in the assembly of the budworm microbiome. While spatial location and spatial structure were also important drivers of spruce budworm bacterial community composition, none of the environmental variables we measured could explain the variation among sites, and identifying the drivers of this spatial variation remains an open question that will need to be addressed by future studies.

1.3 Introduction

Insect microbiomes vary greatly both among species and geographically. Despite the numerous beneficial relationships having been found between insects and their microbiota (Breznak, 1982; Calderón-Cortés et al., 2011; A. H. Smith et al., 2015; Warnecke et al., 2007) knowledge of the processes governing the biogeography and assembly of insect microbiomes is lacking (Engel & Moran, 2013). Given the ecological importance of insects (Dimopoulos, 2003; Hung et al., 2018; Wright, 1986) and the importance of microbe-associated benefits on hosts and ecosystems, there is increasing interest in understanding the biogeography of insect microbiomes. Some studies have suggested that Lepidoptera may lack a persistent, resident gut bacterial community and that their microbiomes are instead populated mostly by transient bacteria that are derived from dietary sources (Hammer et al., 2017; Schwarz et al., 2018). One such important Lepidopteran forest pest is the eastern spruce budworm, *Choristoneura fumiferana*, which is one of the most destructive forest insects in North America (Royama, 1984) and is responsible for the defoliation of almost 13,500,000 hectares in the current outbreak in Québec (Ministère des Forêts, 2020).

The spruce budworm is part of a complex, multitrophic food web in which many other organisms affect its population and outbreak dynamics (Royama, 1984). The role of multi-trophic interactions in spruce budworm outbreaks have been explored in detail for insect parasitoids (Eveleigh et al., 2007; Marrec et al., 2018) and avian predators (Crawford & Jennings, 1989; Germain et al., 2021). However, less attention has been given to the role of the spruce budworm microbiome, even though bacteria and fungi could play an important role in budworm population dynamics (Chen et al., 2016, 2020; Xia et al., 2013). In general, little is known about the factors that drive biogeographic variation in microbiome communities and how they respond to environmental gradients (Cohen et al., 2020), and this is particularly true for the spruce budworm (Landry et al., 2015). Improved understanding of what drives

biogeographic patterns in spruce budworm associated bacterial communities could lead to a better understanding of the drivers of budworm spatial population dynamics, which could be used to better understand outbreak mechanics (Broderick et al., 2009).

Biogeographic variation in budworm bacterial communities is influenced by the community of bacteria present in the insect's diet and those that are present in the insect's environment (Nemergut et al., 2013). The community present in the environment are filtered by the budworm host (Engel & Moran, 2013). Consequently, the pool of environmentally available microbes that colonize the budworm microbiome are likely to be determined by a variety of factors including environmental conditions, foliar bacterial communities, host tree species, and other underlying spatial structure. Environmental conditions, including climate, affect the structure of ecological communities (Hernández-Gómez, 2020; Sepulveda & Moeller, 2020) by acting as filters that allow only a subset of the regional species pool to disperse and establish in a habitat (Kraft et al., 2015). Ecological filters are a form of selection, that is, an ecological process by which organisms possessing specific traits are favored in an environment (Nemergut et al., 2013; Vellend, 2016). Abiotic factors such as moisture, temperature, and precipitation all affect bacterial growth (Balser et al., 2010). Climatic differences between sites can therefore modulate the selection and establishment of microbial populations in natural environments. The spatial configuration of populations of microbiome hosts can also affect microbiome community composition by affecting dispersal from a regional species pool (Nemergut et al., 2013; Vellend, 2016). Finally, spatial structure among sites can be used as a proxy for unmeasured spatial or environmental variation (Borcard et al., 2011) that may influence microbiome assembly, filtering, or dispersal. Currently we do not have a strong understanding of the relative importance of these different factors in determining community assembly and biogeography of insect-associated microbiomes.

An ongoing debate in microbial ecology has focused on whether animal-associated microbiomes are composed primarily of resident versus transient taxa (Astudillo-García et al., 2017; Hammer et al., 2019; Martinson et al., 2019). In insects, some studies have demonstrated that some microbes from the diet can establish and grow in the arthropod microbiome (Kennedy et al., 2020). Conversely, Hammer et al. (2017) argued that the Lepidopteran gut was devoid of a resident microbiome, that most bacteria are transient within the insect, and that most bacteria are imported through diet. Previous work has shown that the bacterial communities residing on tree foliage vary spatially and among tree species (Laforest-Lapointe et al., 2016). If environmental and spatial variability impact foliar bacterial communities, and foliage is the spruce budworm's major dietary source of bacterial inoculants, foliar bacterial communities absorbed through foliage consumption should be the main determinant of the budworm bacterial community. Because the spruce budworm feeds on both spruce and balsam fir, bacterial community composition might also vary as a function of host tree species. The budworm itself could also act as an ecological filter and select for only certain bacterial taxa, in which case larval microbiomes could represent a subset of the foliage communities and specific bacterial taxa on the foliage could be consistently excluded by budworm body conditions.

In this study, we investigated what determines spatial variation in spruce budworm bacterial community composition. Our specific objective was to quantify the relative importance of four factors: 1) environmental conditions, 2) foliage bacterial communities, 3) host tree species and 4) spatial structure. We also sought to determine the degree to which these multiple factors interact to shape bacterial community composition. Because Lepidoptera are thought to lack a resident microbiome (Hammer et al., 2017), we hypothesized that the bacterial communities found on the foliage on which budworm feed should reflect the bacterial community found both in and on spruce budworm larvae. We also hypothesized that because foliar bacterial communities vary geographically and among host tree species

(Laforest-Lapointe et al., 2016), spatial structure and host tree species will be important drivers of microbial community structure in both budworm and foliage.

1.4 Methods

1.4.1 Study area

We sampled spruce budworm bacterial communities and associated foliar bacterial communities at 26 sites along a gradient of outbreak density from Gaspésie to the Côte-Nord in Québec to Newfoundland & Labrador (Figure 1.1) in an area spanning more than 25,000 km^2. Our study area included sites near Baie-Comeau, Québec, which is the epicenter of the current budworm outbreak (Ministère des Forêts, 2020), as well as surrounding areas that had lower budworm population densities at the time of sampling. All sites (Supplementary Table 1.1) were forested and contained both balsam fir (*Abies balsamea*) and black spruce (*Picea mariana*). For each site, we calculated mean annual temperature, total annual precipitation, solar radiation, evaporation, climatic moisture deficit, and relative humidity using ClimateNA v5.10 (Wang et al., 2016).

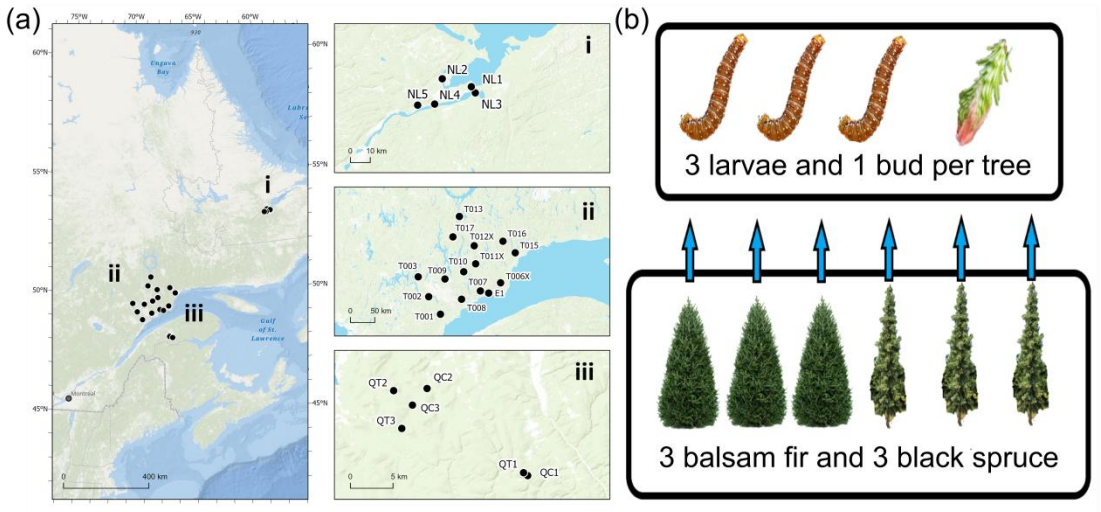

Figure 1.1: A) Map of sampling sites in Québec, Canada and surrounding provinces. Projection: NAD 1983 Québec Lambert Conformal. Sources: ESRI Basemap, MERN 2010. Author: CFR, 2021. B) Schematic of the sampling design at each site.

1.4.2 Sample collection

In the summer of 2015, we collected sixth instar larvae (the last larval stage of the spruce budworm) from all sites. At each site we collected three larvae from three black spruce and three balsam fir randomly selected trees using aseptic techniques. We also collected a foliage sample consisting in a bud from each tree. Tree-level defoliation was visually estimated using the Fettes defoliation index (Fettes, 1951) to be used as an environmental factor. Larval and foliar samples were placed in sterile containers with RNAlater (Thermo Fisher, 2010) and stored on ice for approximately two days before being shipped to the lab, where they were stored at -80^{o}C.

1.4.3 Processing

We characterized the bacterial communities of whole insects at the level of individual trees. Collected larvae from each tree were pooled and homogenized together without any surface sterilization. Foliage samples were crushed by intensive bead-beating with 2.3 mm sterile steel beads in a BioSpec Mini-Beadbeater-16 prior to DNA extraction. DNA was extracted from larval and foliar samples using the Qiagen PowerSoil™ DNA extraction kit according to the manufacturer's instructions. Following DNA extraction, we amplified the bacterial 16S rRNA gene regions V5-V6 using the 799F-1193R primer pair (Chelius & Triplett, 2001) with barcodes and Illumina adapters. Amplified larval and foliar samples were concentration-equalized with the Invitrogen Sequalprep kit before paired-end sequencing with Illumina MiSeq V3 technology to obtain 2x300-nucleotide reads. Extraction negatives (water processed with the extraction kit) and PCR negatives were also sequenced and incorporated in the analysis.

Raw sequencing reads were demultiplexed with the *idemp* software (https://github.com/yhwu/idemp/blob/master) and then imported into the *DADA2* 1.12.1 package (Callahan et al., 2016) in R 3.6.0 (R Development Core Team, 2013)

for processing. We filtered low-quality reads, merged paired-end reads, and trimmed 20 nucleotides from the 5' end of each read to remove barcodes and primers. We then clustered reads into amplicon sequence variants (ASVs) using the *DADA2* default parameters for all functions, and we annotated the taxonomic identity of each ASV by comparison with the SILVA database version 128 (Quast et al., 2012).

1.4.4 Statistical analyses

We first used rarefaction curves to determine if sequencing depth and sampling effort were adequate to capture a sufficient portion of bacterial diversity (Gotelli & Colwell, 2001). Samples with fewer than 4,000 sequences were discarded, and the remaining samples were rarefied to 4,000 sequences per sample; these criteria were sufficient to capture the majority of bacterial ASV diversity while retaining 95% of the samples (Suppl. Figure 1.2). All subsequent analyses except for differential abundance analysis were performed on the rarefied data.

1.4.5 Characterizing spatial structure

To include spatial predictors that account for unmeasured spatial variation, we quantified the spatial structure of our sampling sites using Moran's eigenvector maps (MEMs) (Borcard et al., 2011). MEMs are orthogonal vectors that quantify spatial structure in geospatial data (Legendre et al., 2015). These synthetic variables, referred to as MEM eigenvectors, capture spatial structure at different spatial scales and can be used as predictors in constrained ordination models (e.g., redundancy analysis [RDA]). Use of MEMs in an RDA framework to quantify spatial structure in ecological and genetic data is preferred relative to other methods often used in ecology, such as the Mantel test, due to its lower type 1 error rate (Harmon & Glor, 2010; Legendre et al., 2015; Lisboa et al., 2014).

1.4.6 Larval and foliar communitiy variation

We tested whether bacterial community composition varied significantly among geographic location and between host tree species (Anderson, 2017) using hierarchical modeling with permutational multivariate analyses of variance (PERMANOVAs) performed with the *adonis* function in *vegan* 2.5 in R. Bacterial taxa that were differentially abundant between spruce-feeding and fir-feeding budworm were then identified using the Analysis of Compositions of Microbiomes with Bias Correction (ANCOM-BC) method which models absolute abundance in a linear regression framework (Lin & Peddada, 2020). We used ANCOM-BC with default parameters, with two exceptions: 1) we filtered for taxa present in at least two samples, 2) we enabled structural zeros detection. We then used Principal Coordinates Analysis (PCoA) based on the Bray-Curtis dissimilarity between samples to visualize differences in community composition (Gower, 1966). To determine whether our measured environmental and spatial variables explained overall variation in bacterial community composition, we used the *envfit* function of the R package *vegan* 2.5 (Oksanen et al., 2016) to identify and plot predictors that were significantly correlated with the ordination axes ($p < 0.05$).

1.4.7 Correlation between foliar and larval bacterial communities

The data used for comparing foliar and larval communities included all larval samples that could be associated their homologous foliage sample with enough reads (>4000). We measured the relationship between foliar and larval bacterial communities using the Procrustes rotation test (Mardia et al., 1979; Peres-Neto & Jackson, 2001) using the PCoA-derived ordination solutions for both data sets as inputs. ANCOM-BC was used as before to identify taxa whose abundance was significantly different between larval and foliar samples.

1.4.8 Modeling variation in bacterial community composition

We divided potential predictors of spruce budworm bacterial community composition into four categories to investigate the contribution of each predictor to community variation among samples, while controlling for the effect of the others. The four categories were: 1) environment, which included host tree damage (Fettes defoliation index), elevation and site-based climate variables; 2) foliar bacterial communities; 3) host tree species; and 4) spatial structure which was represented by MEMs. For this analysis, the bacterial community composition of foliage was represented using the scores of the first two axes of the PCoA ordination of the foliage bacterial communities. All continuous predictors were standardized to zero mean and unit variance prior to analysis. The response matrix, which in this case was the matrix of bacterial community relative abundance data, was Hellinger-transformed (Legendre & Gallagher, 2001).

We next identified the best predictors within each of the four categories described above by fitting models of community structure using redundancy analysis (RDA). For each category, we constructed an RDA model predicting bacterial community composition, and we reduced the set of potential predictors in the category using a bidirectional stepwise model selection procedure implemented with the *ordistep* function in *vegan*. Permutation *p*-values for each variable were used as the criterion for variable retention and model simplification. The variables that were retained in each of the four components after model simplification were: 1) precipitation; 2) the scores of the first two axes of the foliar community PCoA ordination; 3) host tree species (i.e., fir or spruce); and 4) two of the eight eigenvectors obtained from the MEMs {1, 8}.

For our final model, we used the same transformed matrix indicated above as the response and the significant variables retained in each of the four categories as predictors. The unique and shared contributions of each predictor category to overall

biogeographic variation in bacterial community structure among budworm larvae was then assessed using variance partitioning (Borcard et al., 1992). It is important to note that predictors that were excluded from this final analysis may still potentially impact budworm communities; they were removed during stepwise selection on the basis of permutational p-values, but this does not imply that they are necessarily unimportant. Correlations among predictors are shown in Supplementary Tables 1.4 and 1.5.

1.5 Results

1.5.1 Larval and foliar community variation

Spruce budworm bacterial communities were dominated by the phyla *Firmicutes* (39.6±4.0% relative abundance) and *Proteobacteria* (36.3±3.6%), whereas foliage communities were mostly dominated by *Proteobacteria* (53.8±1.5%). We detected a total of 131 ASVs in larvae and 115 ASVs in foliage. Because negative control samples were either very different from experimental samples or had very few reads, indicating that reads in experimental samples represented budworm or foliage bacteria, they were removed from the analysis. The cut-off of 4,000 reads per larval and foliar samples was sufficient to ensure that all ASVs represented all bacterial diversity in our samples, as shown by the asymptotic rarefaction curve (Suppl. Figure 1.1). The average (±SD) number of ASVs per sample was 24 ± 11 for larvae and 28 ± 11 for foliage. There was no significant difference in the average number of ASVs found in larval samples versus foliage samples (ANOVA; $F = 2.437$, $p=0.121$).

Region and sample site explained almost one-third of the variation in larval bacterial community composition (PERMANOVA; Region: $R^2 = 0.04$, $p = 0.001$, Site: $R^2 = 0.25$, $p = 0.001$), indicating that insects at each site had a distinctive bacterial community structure (Table 1.1) and that most of spatial variation was between individual sites instead of between regions. Host tree species nested within sample

site explained an additional 21% of the variation in community structure (PERMANOVA R^2 = 0.21, p = 0.026). At deeper phylogenetic levels like Genus and Family, the variation explained by site and region stayed the same while host tree species became a unsignificant term in the model (Table 1.1). Differential abundance analysis detected ASVs that were positively associated with larvae that fed on balsam fir, which were mostly *Firmicutes* and *Proteobacteria*, specifically *Acetinobacter* and *Pseudomonas* genera (Figure 1.2). When we overlaid environmental vectors on ordinations of community composition (Figure 1.4), we found that both climate factors and MEMs 1, 5 and 8 were significantly correlated with both ordination axes. When looking at which spatial patterns these MEMs represented, MEM1 appeared to capture the spatial variation among the three regions, while MEM 5 and 8 captured uncharacterized fine-scale spatial variation (Suppl. Figure 1.2).

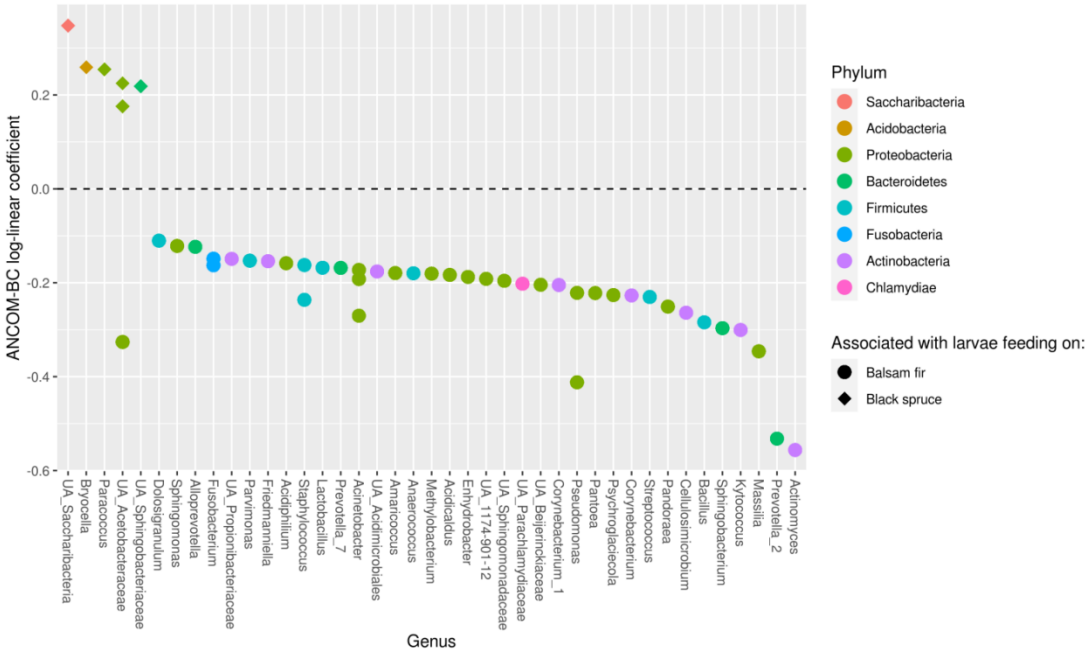

Figure 1.2: ANCOM-BC discriminant analysis of differentially abundant ASVs in spruce budworm larvae coming from different host tree species. Most differentially abundant ASVs were positively associated with balsam fir and were from the *Proteobacteria* and *Actinobacteria* phyla.

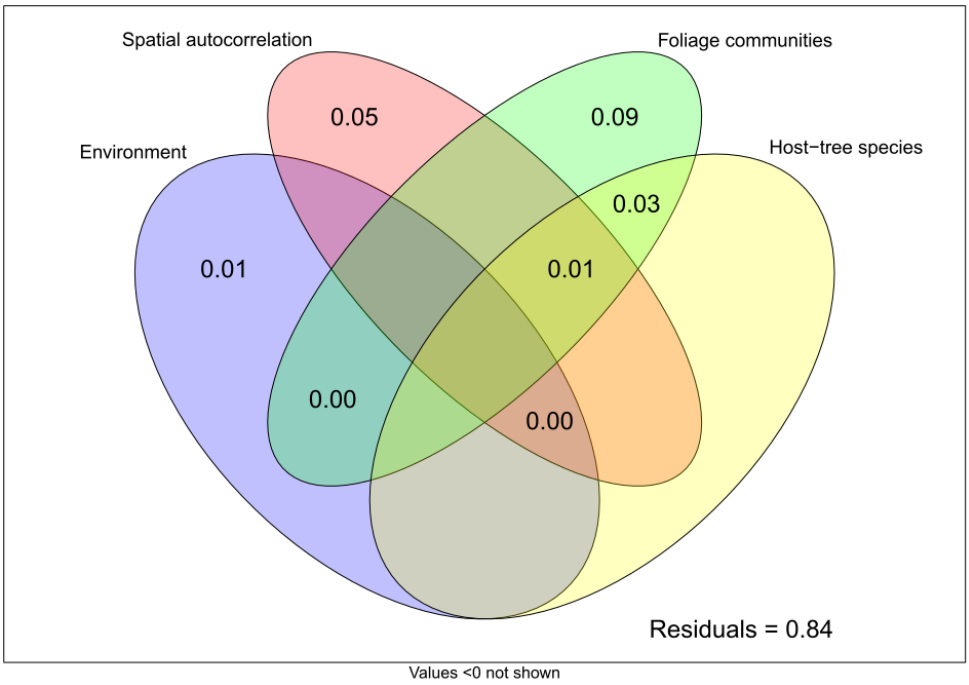

Figure 1.3: Variance partitioning analysis dividing variation in spruce budworm bacterial community structure among four possible predictors: environmental variables, foliage bacterial communities, host tree species and spatial structure (Moran's eigenvector maps (MEMs)). Values in the figure represent the proportion of variation in bacterial community structure explained by each set of variables and were determined based on a partial redundancy analysis of bacterial community composition.

Table 1.1: Variation in spruce budworm bacterial communities as a function of sampling site and host tree species, tested using a Bray-Curtis dissimilarity-based PERMANOVA on ASV, genus, and family-level community tables. Host tree species was nested within site which was nested within region for this analysis.

ASV						
	Df	**Sum Sq**	**Mean Sq**	**F-Value**	**R^2**	**P-Value**
Region	**2**	**1.913**	**0.95629**	**2.5538**	**0.04435**	**0.001**
Site (within region)	**23**	**10.703**	**0.46535**	**1.2427**	**0.24821**	**0.001**
Tree species (within site)	**22**	**9.162**	**0.41644**	**1.1121**	**0.21246**	**0.026**
Residuals	57	21.344	0.37446		0.49498	
Total	104	43.122			1	
Genus						
	Df	**Sum Sq**	**Mean Sq**	**F-Value**	**R^2**	**P-Value**
Region	**2**	**2.116**	**1.01818**	**3.6888**	**0.06254**	**0.001**
Site (within region)	**23**	**8.53**	**0.37089**	**1.2929**	**0.2521**	**0.002**
Tree species (within site)	22	6.84	0.31089	1.0838	0.20213	0.202
Residuals	57	16.351	0.28686		0.48322	
Total	104	33.837			1	
Family						
	Df	**Sum Sq**	**Mean Sq**	**F-Value**	**R^2**	**P-Value**
Region	**2**	**2.1575**	**1.07876**	**4.2325**	**0.07016**	**0.001**
Site (within region)	**23**	**8.0207**	**0.34872**	**1.3682**	**0.26084**	**0.003**
Tree species (within site)	22	6.044	0.27473	1.0779	0.19655	0.231
Residuals	57	14.5277	0.25487		0.47245	
Total	104	30.7499			1	

1.5.2 Modeling variation in bacterial community structure

We used variance partitioning to quantify the unique and shared contributions of environmental variation, foliar bacterial communities, host tree species, and spatial structure to budworm bacterial community composition. Collectively, all four components explained 16% of the overall variation in spruce budworm bacterial community composition (Figure 1.3, Suppl. Table 1.6). Of the four components, spatial structure and foliage bacterial communities best explained variation in spruce budworm larval microbiome composition among sites; spatial structure explained a unique 5% of this variation and foliage bacterial communities explained another unique 9%. Host tree species had shared components with spatial structure and foliage communities that explained 1% and 3% of the variation, respectively.

When larval bacterial communities were analyzed separately for the two host tree species, we found that different environmental factors were correlated with the first

two ordination axes for each host tree (Suppl. Figure 1.3). Spatial structure (MEMs) and environmental factors including precipitation, evaporation, temperature, moisture, and solar radiation were more strongly associated with bacterial community structure in larvae feeding on balsam fir than in larvae feeding on black spruce. PCoA axes for bacterial communities found in larvae feeding on black spruce were only correlated with the spatial structure variable MEM5, which captured fine scale spatial variation, elevation and evaporation (Suppl. Table 1.2).

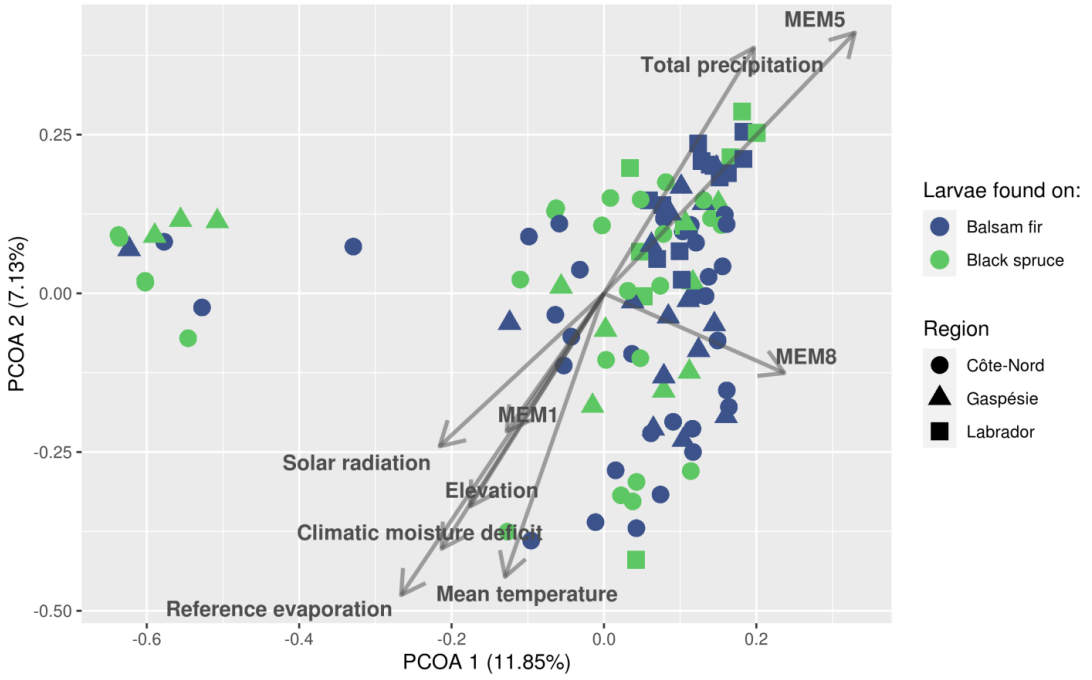

Figure 1.4: Principal coordinates (PCoA) ordination based on the Bray-Curtis dissimilarity between bacterial communities, Vectors indicate variables that were significantly correlated with ordination axes ($p < 0.05$) based on a randomization test. Variance explained by each PCoA axis is shown in axis labels. Most predictor variables clustered into two separate groups of intercorrelated vectors that were correlated with the two ordination axes, indicating that the composition of spruce budworm bacterial communities covaried with several measured environmental variables.

1.5.3 Correlation between foliar and larval bacterial communities

The composition of spruce budworm larval bacterial communities was highly correlated with that of the foliage on which they were feeding (Procrustes test; R^2 = 0.64, p = 0.001). Differential abundance analysis (Figure 1.5) showed that few taxa were differentially abundant between larval and foliage communities: *Actinobacteria* were more abundant in larval samples and *Proteobacteria* were more abundant in foliar samples (Figure 1.5). Like larval communities, variation in foliar bacterial communities was significantly associated with geographic location (i.e., site), which explained 71% of the variance among samples (PERMANOVA; R^2 = 0.71, p = 0.044).

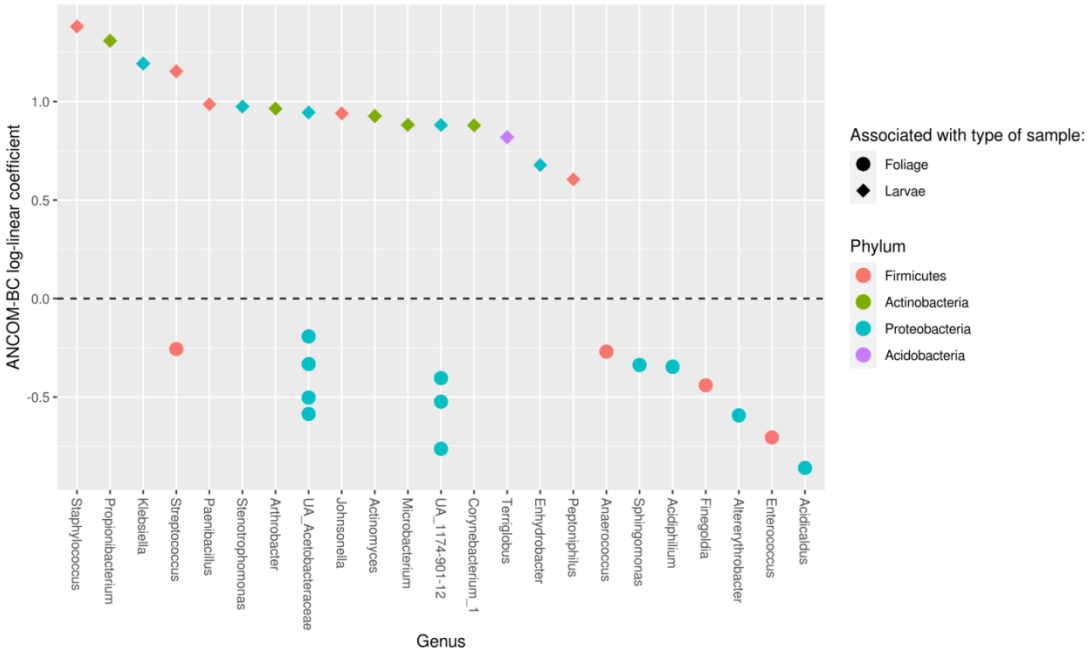

Figure 1.5: ANCOM-BC discriminant analysis of differentially abundant ASVs between larvae and foliage they were feeding on. Most differentially abundant ASVs were from the *Proteobacteria* and *Firmicutes* phyla.

1.6 Discussion

We assessed how environmental factors, host tree species, foliar communities and spatial structure affected the biogeography of microbes inhabiting the spruce budworm. We accomplished this by characterizing the bacterial communities of budworm larvae collected on two different host tree species at 26 different sampling sites in eastern Canada. We specifically sought to test the hypothesis that budworm-associated bacterial communities are determined mainly by the bacterial communities on the foliage that they consume. We confirmed this hypothesis in identifying a significant correlation between the ordination solutions for both communities (Procrustes test; $R^2 = 0.64$, $p = 0.001$), and given the proportion of variance in larval

community structure explained by foliar communities (Figure 1.3). We also confirmed our hypothesis that host tree species and spatial structure affect budworm bacterial communities.

We identified significant biogeographic variation in spruce budworm bacterial community structure among the 26 sampling sites spanning more than 800 kilometers. Most of this variation was among sites (25%) and between different host tree species within sites (22%; Table 1.1). A small portion of the variation was also explained by region (4%) showing that most of the spatial variation in budworm communities was present at a finer spatial scale than region. Among-site variation could be linked to uncharacterized fine scale spatial variation such as unmeasured environmental variation or further variation among foliar bacterial communities. Indeed, we identified a significant contribution of fine scale MEMs that could reflect these hypothesized processes. Previous work has also found that foliar bacterial community structure is highly variable among sites (Laforest-Lapointe et al., 2016).

Strong among-site differences in community composition suggest that dispersal limitation or filtering by spatially autocorrelated environmental factors are driving budworm bacterial community composition. Despite these site-level differences, none of the environmental or climatic variables we measured explained this variation (Figure 1.3). Instead, MEMs that capture spatial structure, but do not identify its root cause, were among the best predictors of spruce budworm bacterial community structure. As a result, the biological factors underlying this spatial variation in community structure remain unknown. The fine and coarse scale MEMs that explained community structure likely capture unmeasured climatic factors or dispersal limitation of the budworm, their associated bacterial communities, or both. Previous studies have shown that climatic conditions have a large influence on moth dispersal (Halsch et al., 2021) and could impact budworm distribution (Régnière et al., 2012). However, because of the similarity between foliar and larval communities, the

spatial variation we measured is likely to also be influenced by free-living bacteria dispersal captured by larval communities in addition to bacteria being transported by budworm larvae. Studies have shown that ASV-based bacterial analyses can overestimate the importance of dispersal limitation (Raulo et al., 2021; Stothart et al., 2021). For this reason, we repeated our analyses of community variation explained by sampling site at different phylogenetic levels (Table 1.1). The results were similar, indicating that dispersal limitation is relatively constant whether ASVs are aggregated at the family or genus level and that the dispersal traits of budworm-associated bacterial communities could be conserved at deeper phylogenetic level (Martiny et al., 2013; Martiny et al., 2015).

Host tree species had a measurable impact on both larval and foliar bacterial communities. More than 22% of the overall variance in bacterial community structure was explained by the identity of host tree species. With respect to the determinants of these differences, we found that budworm bacterial communities from different host tree species were correlated with different explanatory variables; for example, most environmental conditions had an impact on the bacterial communities of larvae on balsam fir but not on those on black spruce (Suppl. Table 1.3). The effect of host tree species on spruce budworm larval bacterial communities could be explained by differences in the chemical composition of the foliage of the two host tree species and which would drive differences in foliage-associated bacteria that are then ingested by the insect. Animals who are a source of bacterial dispersal on foliage could also prefer one tree species over another (Beltrán & Wunderle, 2013), impacting the available dispersal pool to the budworm on different host tree species. Additionally, conifer needles have been shown to generally inhibit bacterial growth in the spruce budworm due to the antibacterial effect of tannins (Pang, 2010), and this effect could vary between tree species.

Overall we found that microbial communities on the budworm's foliage diet were the most important driver of budworm microbiome composition. This result suggests that the key process determining bacterial community structure within the spruce budworm is colonization by microbes from their diet, and is consistent with the hypothesis that Lepidoptera lack a resident microbiome (Hammer et al., 2017; Schwarz et al., 2018), which predicts that larvae should have the same bacterial community as that found on their foliage diet. Our demonstration of the broad similarities between the microbiomes of the spruce budworm and its foliage diet provides some of the first field-based evidence to support the hypothesis that some Lepidoptera lack a persistent resident microbiome. However, there were some differences in taxonomic composition between the spruce budworm and foliage microbiota and some taxa that were associated with the budworm microbiome, and these differences likely represent taxa that are able to establish and grow in the budworm microbiome, as has been observed in other arthropod-associated microbial communities (Kennedy et al., 2020); for example, *Actinobacteria* were significantly more abundant in larvae as were several ASVs from diverse phyla, whereas *Proteobacteria* were more abundant in foliage. The *Actinobacteria* and other taxa associated with larvae but not foliage may originate elsewhere in the larvae's environment, or from vertical transmission. These taxa could also be selectively enriched from species in the foliage-associated communities that were too rare to be detected, due to selection by the biotic or abiotic conditions in the budworm microbiome leading to strong increases in abundance and increased possibility of detection. Thus, while the majority of the diversity in the budworm microbiome is likely transient taxa originating from foliage, some bacterial taxa do appear to be likely residents of the microbiome. Foliar communities are themselves impacted by environmental factors including air quality and climate (Espenshade et al., 2019; Laforest-Lapointe et al., 2016; Magan & McLeod, 1991), suggesting that future studies that take into account direct and indirect environmental drivers of animal

microbiomes will be required to understand the full spectrum of factors driving the biogeography of these microbiomes.

The spatial and environmental factors we measured explained almost a sixth (16%) of variance in community composition, with the variance explained by foliage communities peaking at 9%. The poor explanatory power of our measured variables could be due to the importance of stochastic processes in budworm microbiome community assembly and biogeography, such as dispersal of insect hosts and microbes and subsequent colonization by microbes. The scale at which we quantified spatial structure may not capture well such dispersal and colonization processes. Genetic variation in budworms or host trees could be a potential explanation for the important amount of spatial variation in budworm bacterial communities yet to be explained. Spruce budworm spatial genetic structure in our study area (Côte-Nord and Gaspésie, Figure 1.1) was generally homogenous, with some subtle genetic differences between the two regions (Larroque et al., 2020). Because most of the inter-site bacterial community variation was at a finer spatial scale than that of these regional budworm differences (Table 1.1), it seems likely that the impact of budworm genetic variation on bacterial communities is minimal. However, as foliar communities are closely linked to larval communities, linking host tree genetic variation in the outbreak area to budworm bacterial communities could be a promising avenue of future research.

1.7 Conclusions

We have shown that the community assembly and biogeography of spruce budworm bacterial communities is shaped by spatial structure among sites, by foliage communities, and by host tree species. Most of the variation in budworm communities was mirrored in the communities of foliage they ate, showing that foliage bacterial communities are likely the main determinant of budworm

microbiome composition. While spatial location (i.e., site) and spatial structure of sites was an important indicator of spruce budworm bacterial community composition, none of the environmental variables we measured could explain the variation among sites, and identifying the drivers of this spatial variation remains an open question that will need to be addressed by future studies. We also determined that host tree species is an important determinant of both spruce budworm microbiome composition and the sensitivity of that microbiome to the environment, as the microbiomes of budworms living on balsam fir were more strongly correlated with other spatial and environmental predictors. Finally, we found support for the hypothesis that Lepidopteran microbiomes are mostly non-resident and likely originate from the foliage diet. However, subtle differences in microbial communities between larvae and foliage suggest that some bacteria do establish and grow in the budworm microbiome, and that dispersal of bacteria from other sources than foliage and differences in environmental filtering between larval bodies and foliage play a role in the assembly of the budworm microbiome. Our results advance understanding of the factors that drive the assembly and biogeography of bacterial communities associated with an ecologically and economically important forest insect, and open the door to understanding how these factors impact the biogeography of the budworm itself and its outbreak consequences.

1.8 Acknowledgements

We thank Rob Johns, Paul Whalen, Jérémie Brisson and Marie-Line Buteau for contributing to sampling. We also thank Geneviève Bourret for assistance with sample processing and sequencing. We thank the reviewers for their input in improving this manuscript.

CHAPITRE II

EASTERN SPRUCE BUDWORM BACTERIAL COMMUNITIES DIFFER BETWEEN SEXES AND BETWEEN GUT TISSUES AND WHOLE LARVAE

Cet article est en préparation.

Mathieu Landry, premier auteur.
Département des sciences biologiques, Université du Québec à Montréal,
141 Av. du Président-Kennedy, Montréal, Québec, H2X 1Y4
landry.mathieu.5@courrier.uqam.ca

Patrick M.A. James
Institute of Forestry and Conservation, University of Toronto,
33 Willcocks St, Toronto, Ontario, M5S 3B2
patrick.james@utoronto.ca

Daniel Kneeshaw
Département des sciences biologiques, Université du Québec à Montréal,
141 Av. du Président-Kennedy, Montréal, Québec, H2X 1Y4
kneeshaw.daniel@uqam.ca

Steven W. Kembel
Département des sciences biologiques, Université du Québec à Montréal,
141 Av. du Président-Kennedy, Montréal, Québec, H2X 1Y4
kembel.steven_w@uqam.ca

2.1 Résumé

Bien que notre compréhension de l'impact de différents facteurs spatiaux et environnementaux sur l'assemblage du microbiome chez les insectes défoliateurs ait évolué, on en sait peu sur la façon dont les attributs morphologiques et physiologiques de l'hôte affectent les communautés microbiennes associées aux insectes. En particulier, on en sait peu sur la façon dont la structure des communauté microbiennes varie entre les sexes d'insectes ou à travers différents compartiments morphologiques de l'insecte hôte, tels que les tissus intestinaux par rapport à la larve dans son ensemble. Pour pallier ce manque de connaissances, nous avons caractérisé les communautés bactériennes associées à la tordeuse des bourgeons de l'épinette (TBE), *Choristoneura fumiferana*, l'un des ravageurs les plus destructeurs d'Amérique qui se nourrit des populations de sapin et d'épinette dans le nord-est des États-Unis et du Canada. Nous avons utilisé le séquençage du gène de l'ARNr bactérien 16S des communautés bactériennes associées à la TBE de plusieurs sites dans l'est du Canada pour quantifier les différences de composition et de diversité entre les tissus intestinaux des larves mâles et femelles ainsi qu'entre les communautés de tissus intestinaux et les larves entières. Nous avons également mesuré comment l'emplacement géographique interagissait avec le sexe des larves et la composition des communautés bactériennes associées à la TBE. Nous avons constaté que les tissus intestinaux de la TBE présentaient des différences mesurables mais subtiles dans la composition de la communauté entre les sexes en termes de diversité de la communauté et d'abondance de taxons spécifiques. Les tissus intestinaux de la TBE contenaient des communautés significativement moins diversifiées que les larves entières et présentaient moins de variation dans leur composition que les communautés de larves entières. L'emplacement géographique avait un impact plus important sur la composition de la communauté des tissus intestinaux que sur celle des larves entières. Nous concluons que les communautés des tissus intestinaux de la TBE sont spatialement plus homogènes que les communautés de larves entières, et que le sexe larvaire a un faible impact sur la structure globale de ces communautés bactériennes. Nos résultats mettent en évidence que les communautés des tissus intestinaux des Lépidoptères ne sont pas représentatives des communautés de l'insecte entier et donc que le choix des structures morphologiques à échantillonner lors de la quantification des microbiomes d'insectes a des impacts majeurs sur la caractérisation de la structure des communautés bactériennes.

2.2 Abstract

While our understanding of how different spatial and environmental factors impact microbiome assembly in defoliating insects has evolved, not much is known about how host morphological and physiological attributes affect insect-associated microbial communities. In particular, little is known about how microbial community structure varies between insect sexes, or across different morphological compartments of the insect host such as gut tissues versus the larva as a whole. To address this lack of knowledge, we characterized the bacterial communities associated with different sexes and organs of the spruce budworm (*Choristoneura fumiferana*), a destructive pests that feeds on fir and spruce trees in North America . We used 16S bacterial rRNA gene sequencing of microbial communities associated with spruce budworm from multiple sites across eastern Canada to quantify bacterial community diversity and composition differences between the gut tissues of male and female larvae as well as the whole larvae. We also measured how geographic location and larval sex influenced community composition. We found that budworm gut tissues showed measurable but subtle differences in community composition between sexes in terms of community diversity and the abundance of specific taxa. Spruce budworm gut tissues contained significantly less diverse communities than whole larvae and showed less variation in their composition than whole larvae communities. Geographic location had a larger impact on gut tissue community composition than on that of larvae. We conclude that spruce budworm larval gut tissue communities are more spatially homogeneous than whole larval communities, and that larval sex has a weak impact on the overall structure of these microbial communities. Our results highlight that Lepidopteran gut tissue communities are not representative of communities of the whole insect and thus that the choice of which morphological features to sample when quantifying insect microbiomes has major impacts on the resulting measures of microbial community structure.

2.3 Introduction

Microbial communities that live in association with animals are extremely diverse and can be functionally crucial to their hosts (Levin et al., 2021). Many consider host-associated microbial communities to function as a separate organ with its own self-regulation and metabolic roles (Baquero & Nombela, 2012). The role of the microbiota in the physiology, metabolism, and behaviour of the host, and what determines the assembly of these communities is less well known in insect systems compared with other animals such as mammals (Engel & Moran, 2013). In particular, the factors that determine inter- and intra-individual variation in microbial composition and diversity in insect microbiomes is still an open question (Engel & Moran, 2013). In previous studies, we have shown that geographical location, host-tree species and foliage bacterial communities are important determinants of bacterial community composition in spruce budworm larvae (Landry et al., 2022). In this study, we focus on quantifying the role of a potentially important source of inter-individual variation in microbiome structure for the economically and ecologically important forest Lepidopteran species eastern spruce budworm (*Choristoneura fumiferana)*, larval sex. We also focus on characterizing intra-individual variation within the larval body between communities of gut tissues versus whole larvae tissues.

Although the effects of sex on insect microbiomes has been studied in mosquitoes and fruit flies (Han et al., 2017; Segata et al., 2016), to date no studies have characterized differences in microbiome composition in Lepidoptera such as the spruce budworm. In a previous study, we identified unexplained differences in gut tissue microbiome composition among evenly sized subsets of spruce budworm larvae, and hypothesized that these unexplained differences in microbiome structure arose due to differences in sex (Landry et al., 2015). While numerous factors including spatial and environmental variation are known to influence microbial community composition in insects (Landry et al., 2022), host sex also has the

potential to play a role (Jašarević et al., 2016). In humans, host sex has been found to affect microbiome composition and function in the gut, and to have important impacts on immunity and overall health (Jašarević et al., 2016; Vemuri et al., 2019). Sexual hormones and other physiological sex differences have been shown to drive differences in the gut microbiome from birth to death and to modulate the impact of diet, antibiotics, and environment on gut microbiome composition (Valeri & Endres, 2021). In other mammals such as mice, microbiome composition has been shown to differ between sexes due to differences in behaviour (Bridgewater et al., 2017). In mosquitoes and fruit flies, reproductive and digestive systems from different sexes harbour different microbial taxa (Han et al., 2017; Segata et al., 2016). Differences in size, oviposition behaviour, parasite resistance and diet among the sexes have been found in multiple insects (Baker et al., 1987; Fanson et al., 2013; Nakashima & Hirose, 2003; Willink & Svensson, 2017), and these differences could be associated with intersex variation in microbiome composition.

Although the insect gut serves as habitat to many species of microbes (Engel & Moran, 2013; Moran et al., 2019), the rest of the insect is also home to an ecologically and functionally important bacterial community (Otti, 2015). While we know that different animal tissues and body parts are home to distinctive microbial communities based on work in humans and other mammal species (Aagaard et al., 2013; Costello et al., 2009; Dekaboruah et al., 2020; Franzosa et al., 2015; Grice et al., 2009; Human Microbiome Project Consortium, 2012; Stehlikova et al., 2019; Turnbaugh et al., 2007; Voigt et al., 2015), it is not known how microbial community structure varies among body parts in insects in general, and for Lepidoptera in particular. As different selective and dispersal processes likely play a role in different anatomical structures of the insect (Vellend, 2016), it would logically follow that different microbe communities are associated with different anatomical structures of insects, as we see in mammals (Dekaboruah et al., 2020; Stehlikova et al., 2019). For example, microbial communities on the cuticle of insects are directly exposed to

fluctuations in temperature and light, and would be susceptible to colonization by microbes dispersing from the insect's environment. In contrast, gut tissue communities are likely influenced by the distinctive environment of the gut and would primarily be colonized via the insect's diet. Addressing the question of how microbial communities vary across different insect body parts is important both in order to understand the ecological and evolutionary processes driving microbial community assembly in different habitats across the insect body, but also because we frequently use measures of easily obtained samples as a proxy for direct measures of whole-body or gut microbiota structure. For example, whole insects are often used as proxies of gut microbiome composition (Ceja-Navarro et al., 2015; Phalnikar et al., 2019). Characterizing differences between gut tissue and whole insect bacterial communities is thus a crucial step both in identifying how methodological approaches related to the choice of which morphological features to sample influence measured insect microbial communities, and in understanding how microbial communities assemble and persist in different compartments of the Lepidopteran body.

In this study, we first examined how spruce budworm gut tissue bacterial communities vary as a function of larval sex. Our response variables consisted of alpha diversity measured using the Shannon index (Shannon, 1948) and beta-diversity measured with the Bray-Curtis dissimilarity (Bray & Curtis, 1957). These variables will be referred to throughout the manuscript as community diversity and variation in composition, respectively. Such differences between gut tissue communities of spruce budworm larvae of different sexes have been previously hypothesized for the spruce budworm (Landry et al. 2015) as well as for other insect species (Han et al., 2017; Segata et al., 2016). Consequently, we predict that gut tissue communities will vary among sexes in terms of community composition and diversity. We also examined differences in microbial community composition and diversity between larval gut tissues and whole larvae, both to understand how microbial community assembly varies between these body parts, and to assess

whether whole larval samples serve as an effective proxy for direct measurement of gut tissue microbial community structure. Here, we hypothesized that greater diversity exists in whole larvae than in gut tissue communities, as the larval cuticle is exposed to the environment and dispersal from the environment, whereas the gut is a more selective environment that is inoculated through feeding. We also hypothesized that gut tissue communities would be more sensitive to the host tree species ingested in their diet (i.e., fir vs spruce). Testing these hypotheses will contribute to our understanding of how bacterial communities associate with their insect host and help generate relevant questions about the importance of these communities for insect health.

2.4 Methods

2.4.1 Study area

We sampled spruce budworm larvae at five forested sites within the current outbreak near Baie-Comeau, Quebec (Figure 2.1). The selected sites dominated by spruce budworm host tree species: balsam fir and black spruce. Balsam fir is the primary host, due to greater phenological synchrony between larval development and bud development (Pureswaran et al. 2015). Black spruce is a secondary host that is less vulnerable to defoliation. Sites spanned broad gradients of latitude, longitude, outbreak density, and elevation (Figure 2.1).

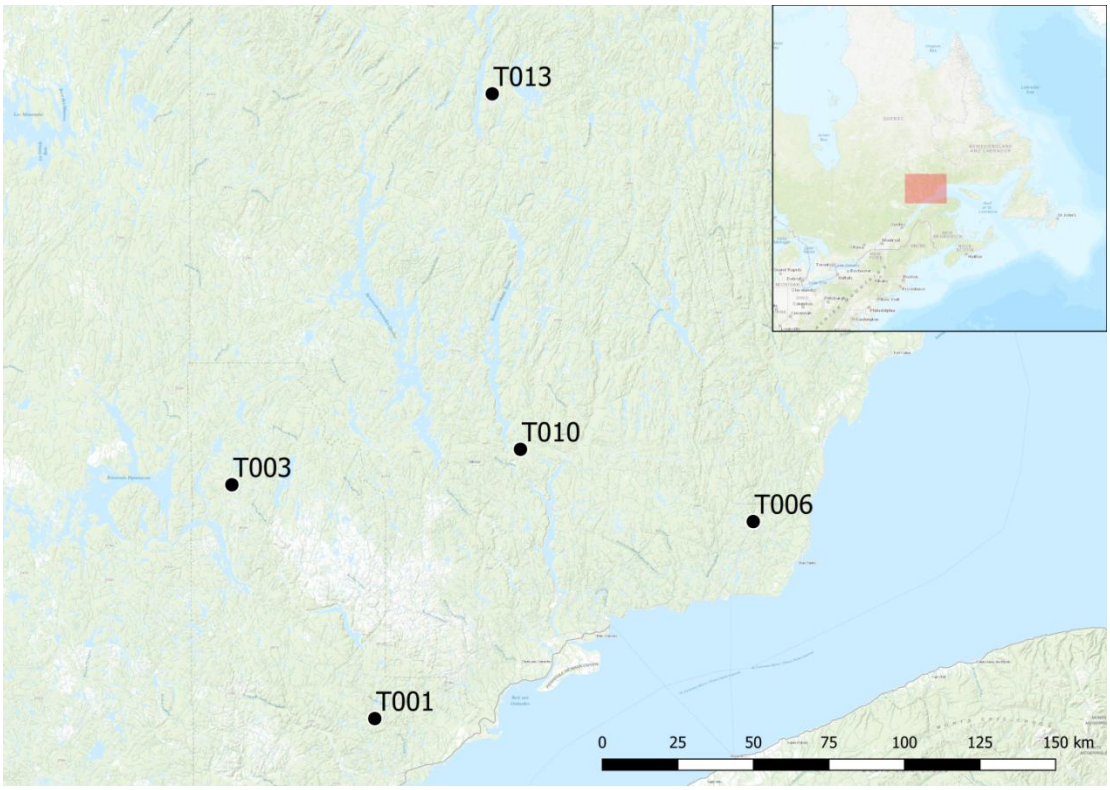

Figure 2.1: Map of sample collection sites in Quebec, Canada.

2.4.2 Sample collection

We collected fifth and sixth instar spruce budworm larvae in the summers of 2015 and 2016. Over its life cycle the spruce budworm passes through six larval stages and the post-winter instar stages (2 to 6) occur over approximately six weeks. The last three stages cause the most damage as larvae are the largest at this time. Six trees (three black spruce and three balsam fir) were randomly selected at each site and three larval samples were collected from each tree using aseptic technique and stored in sterile containers. Samples were stored in RNAlater on ice for approximately 2 days, then shipped to Montreal where they were stored at -80^{o}C. Larvae collected in 2015 were used for whole larvae experimental group whereas larvae from 2016 had

their gut tissues dissected aseptically and were sexed by assessing the presence or absence of male gonads.

2.4.3 Sequencing and sequence data processing

Gut tissues and whole larvae were collected and processed using the Qiagen PowerSoil DNA extraction kit. After DNA extraction, we amplified the *16S* rRNA gene variable regions V5-V6 (799F-1193R primers) (Chelius & Triplett, 2001) and added barcodes and Illumina adapters. Amplified samples were equalized with the Invitrogen Sequalprep kit before paired-end sequencing with Illumina MiSeq V3 technology (paired end 2x300 nucleotide reads). Reads were demultiplexed with the idemp software (https://github.com/yhwu/idemp/blob/master/LICENSE). Extraction (water processed with the extraction kit) and PCR negatives were also sequenced and incorporated in the analysis. For gut tissue and whole samples separately, DNA extraction and PCR were performed in a randomized design where samples were processed independently of their site, host-tree, and sex. We curated, merged, sequenced and trimmed 20 nucleotides from each read to remove barcodes and primers with the DADA2 1.12.1 package (Callahan et al., 2016) in R software 3.6.0 (R Development Core Team, 2013). We then clustered them into amplicon sequence variants (ASVs) using default parameters, removed chimeras according to DADA2's standard operating procedures with the "consensus" method, and assigned taxonomy to ASVs by comparison with the SILVA database version 128 (Quast et al., 2012). Our dataset comprised 3,747,406 reads over 1556 ASVs in 154 samples, for an average of 24,333 reads per samples.

2.4.4 Removal of contaminant sequences

ASVs associated with negative DNA extraction and PCR controls were removed with the *decontam* 1.1.2 package (Davis et al., 2018). Datasets were decontaminated as a function of their respective extraction and PCR controls. Contaminant ASVs

detection was performed with the prevalence method with a very aggressive threshold (0.5) to ensure the identification of control-associated ASVs as contaminants. This meant that ASVs that appeared in more control samples than true samples were identified as contaminants. Overall, 189 ASVs representing 540,854 sequences were removed from the analysis. Negative controls totaled 149,728 reads over 27 controls, for an average of 5,545 reads per control (compared to 24,333 reads per sample for experimental samples). To exclude the possibility of sample biomass affecting their community composition, Spearman's correlation test was used to test for correlations between DNA concentration of gut tissue samples and abundance of two important ASVs in gut tissue communities (Best & Roberts, 1975). No correlation was found (Supplementary Table 2.4).

2.4.5 Statistical analyses

We created two datasets from our samples: one containing all reads for comparison of whole larvae and gut tissues, and another containing only reads coming from gut tissues, for comparison between sexes. The dataset with whole larvae samples contained some samples with fewer reads and was rarefied to 1500 reads per sample, whereas our larval sex dataset was rarefied to 3000 reads per sample. We constructed rarefaction curves to determine if sequencing depth (and sampling effort) was sufficient to capture a significant proportion of the microbial diversity both in spruce budworm gut tissues and whole spruce budworm samples (Supplemental Figures 1 and 2). Shannon's index was calculated for each sample (Shannon, 1948) to evaluate alpha diversity and compared between larvae of different sexes and sample type using the Mann-Whitney-Wilcoxon test.

We then used a generalized linear model to model our non-normal and heteroscedastic Shannon index distribution as a function of larval sex and sampling site (McCullagh & Nelder, 2019). We performed non-metric multidimensional scaling (NMDS) ordination of the Bray-Curtis dissimilarity between each sample to

determine and visualize relative variation between the communities of the samples. We used permutational MANOVAs (PERMANOVAs) to test the contribution of each geographical location and larval sex and sample type to variation inside the community (Anderson, 2001). Differential abundance of taxa between budworm of different sexes or sample types was quantified with DESeq2 (Love et al., 2014) on non-rarefied data from which extremely rare ASVs were removed. This was done by removing ASVs whose total relative abundance was less than 0.01 % of all reads, which for the larval sex and gut tissues versus whole larvae datasets represented ASVs with fewer than 276 reads and 99 reads, respectively. This reduced the number of ASVs in the larval sex dataset from 1214 to 133 and from 1556 to 396 in the gut tissues versus larval dataset, specifically for the DESeq2 analysis. The taxonomy of ASVs which were of particular interest was verified by comparing taxonomic annotations with taxonomic identity using the Basic Local Alignment Search Tool (BLAST) algorithm to compare query sequences to the NCBI's Genbank Nucleotide collection (nr/nt) database as a reference (Altschul et al., 1990). Beta-diversity dispersion tests were used to assess the difference in compositional homogeneity between whole larvae and gut tissue communities (Anderson, 2006).

2.5 Results

2.5.1 Larval sex differences

Spruce budworm larval sex was not significantly associated with gut tissue bacterial community composition (PERMANOVA, p=0.33; Table 2.1). However, we did find that female gut tissue communities tended to contain a greater abundance of bacteria from the phylum *Spirochaeta* (Figure 2.2). Conversely, site was strongly correlated with gut tissue bacterial community composition, accounting for 77% of the variation in composition among samples (PERMANOVA, p=0.001; Table 2.1). We did not identify any significant interaction between site and sex effects on microbial

community structure (PERMANOVA, *p*=0.092; Table 2.1). A NMDS ordination on Bray-Curtis dissimilarities indicated that sites T003 and T006 showed some differences in community composition between larvae of different sex (Figure 2.3), although PERMANOVAs with samples from these sites did not show a significant impact of sex on community composition (Supp. Table 2.1).

Table 2.1: Variation in spruce budworm bacterial communities of gut tissue communities as a function of sampling site and larval sex tested using a Bray-Curtis dissimilarity-based PERMANOVA.

	Df	Sum Sq	Mean Sq	F-Value	R2	P-Value
Site	4	20.44	5.11	73.83	0.777	0.001
Sex	1	0.073	0.07	1.05	0.002	0.333
Site*Sex	4	0.46	0.11	1.66	0.017	0.092
Residuals	77	5.33	0.07		0.203	
Total	86	26.31			1	

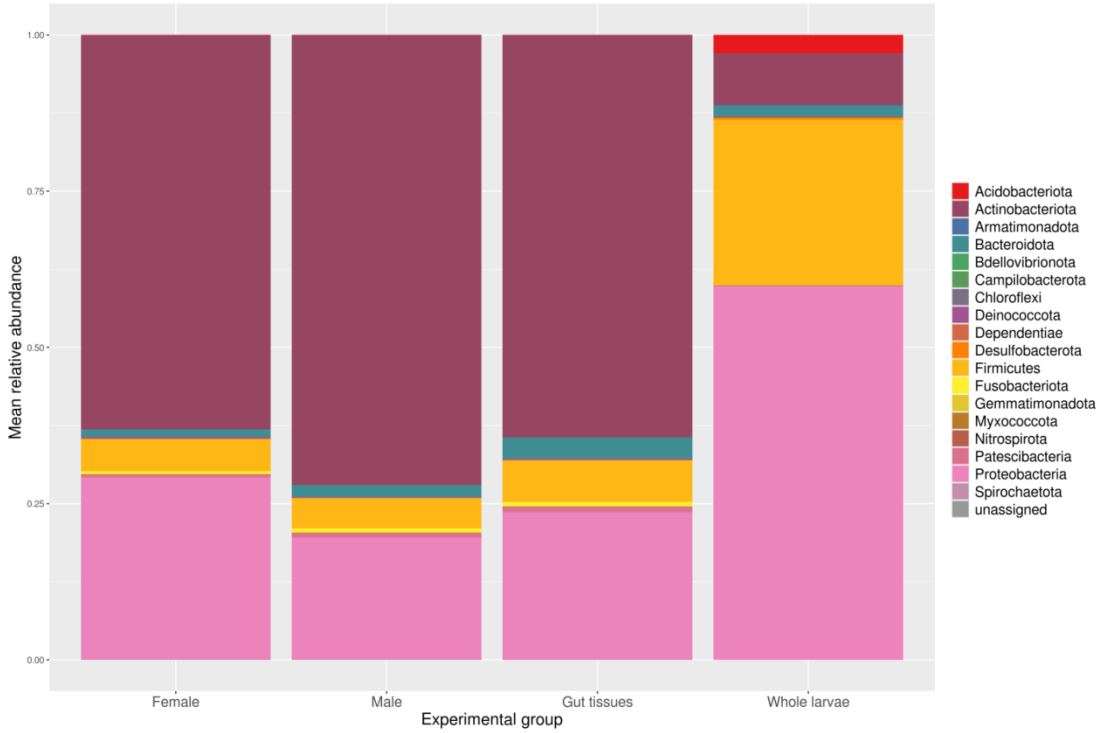

Figure 2.2: Mean relative abundance of all bacterial phyla detected in spruce budworm bacterial communities of gut tissues and whole larvae.

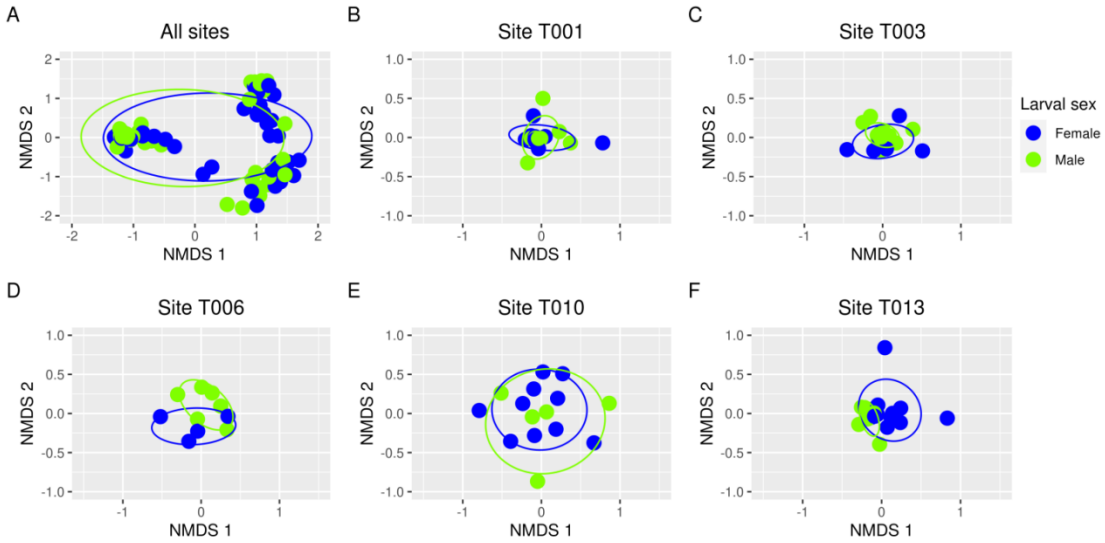

Figure 2.3: NMDS ordinations of spruce budworm gut tissue bacterial communities based on Bray-Curtis dissimilarity for each site, female gut tissue communities are shown in red and male gut tissue communities in blue. 68% confidence interval ellipses are shown overlaid on the ordination for each group.

Contrary to community composition, alpha diversity as measured with the Shannon index differed between larvae of different sexes across all sites (Mann-Whitney-Wilcoxon, p=0.0368; Figure 2.4), females being generally more diverse. Both site and larval sex influenced the Shannon index of gut tissue communities (ANOVA, p<0.001 and p<0.001, respectively; Table 2.2). There was no significant interaction between the effect of site and larval sex on the Shannon index (ANOVA, p=0.73).

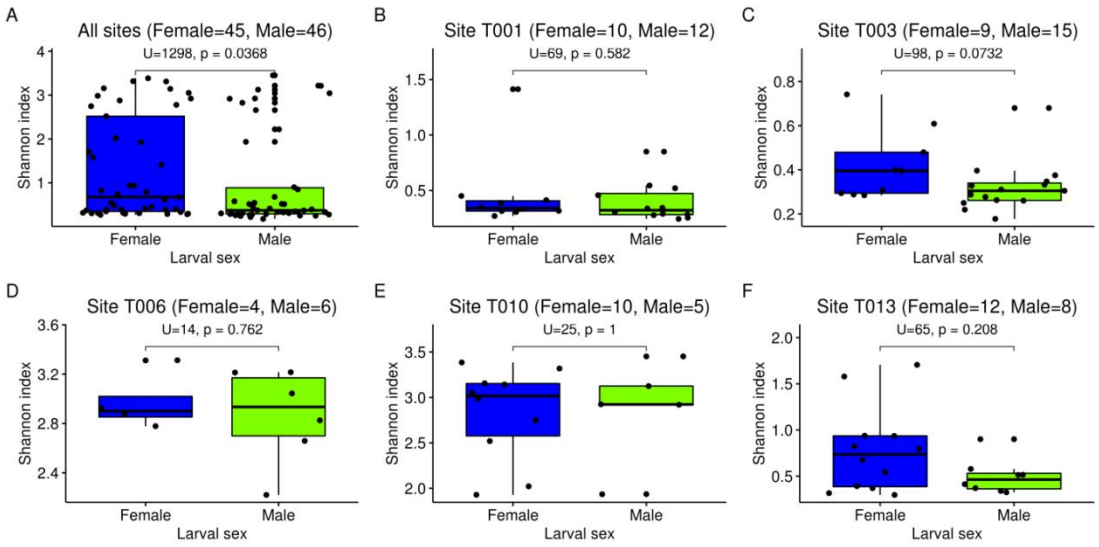

Figure 2.4: Boxplots showing comparisons between bacterial communities of gut tissue samples of male and female spruce budworm larvae in terms of Shannon index of bacterial communities. P-values of Mann-Whitney-Wilcoxon tests between male and female community Shannon index of individual gut tissues are shown over each boxplot.

Table 2.2: Variation in the Shannon index of spruce budworm gut tissue bacterial communities as a function of larval sex and sampling site tested using an ANOVA on a generalized linear model fitted with the "gamma" distribution.

	Df	Sum Sq	Mean Sq	F-Value	P-Value
Sex	1	2.10	2.10	19.22	3.47E-05
Site	4	107.32	26.83	245.92	1.89E-44
Sex*Site	4	0.22	0.06	0.51	0.729
Residuals	81	8.84	0.11		

We detected 27 ASVs that were differentially abundant between male versus female larvae (DeSEQ2 adjusted $p < 0.05$; Figure 2.5). The majority of these differentially abundant ASVs were more likely to be present and more abundant in female samples than in male samples (negative log-fold change values; Figure 2.5). Amongst the most abundant genera associated with male gut tissues were *Pediococcus*,

Capnocytophaga and *Dermabacter* (Supp. Table 2.2). In gut tissues from female larvae the differentially abundant genera included *Gordonia*, *Leptotrichia*, *Staphylococcus*, *Actinomyces* and *Pediococcus*.

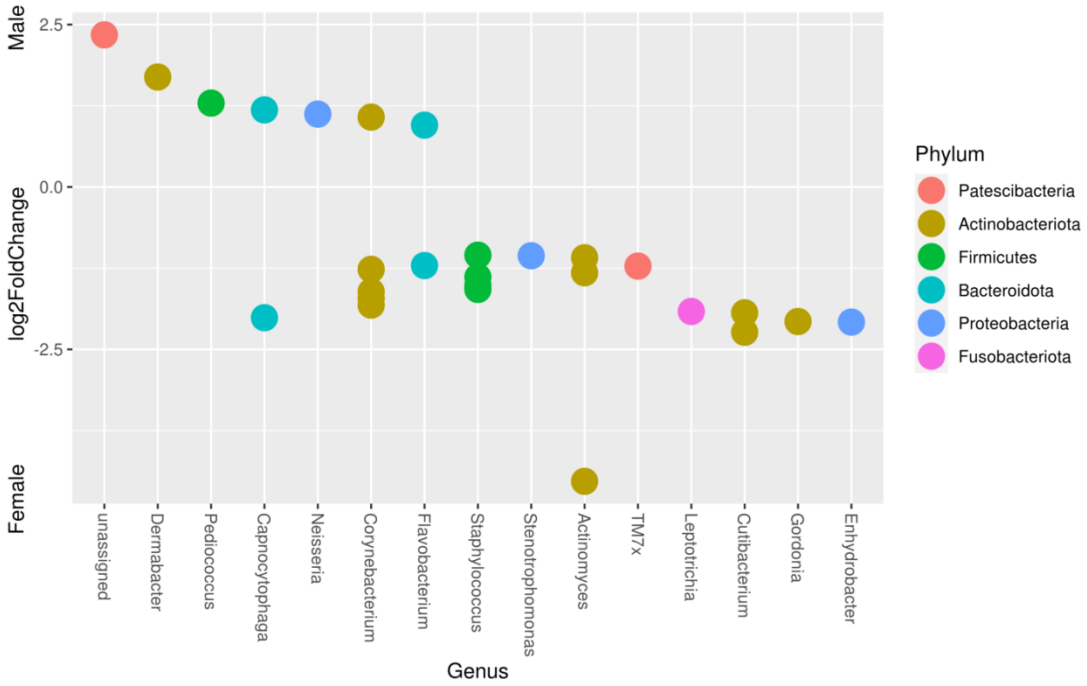

Figure 2.5: DeSEQ2 discriminant analysis of differentially abundant bacterial ASVs in gut tissue bacterial communities of spruce budworm larvae of different sexes. Only ASVs representing more than 0.01% of the entire analysis are shown in the graph.

Two ASVs, ASV1 and ASV5, represented more than 86% of the reads in the gut tissues dataset. They were both unique to this dataset, with zero reads in the whole larvae dataset assigned to these ASVs. ASV1 was identified as *Rhodococcus* sp. and ASV5 as *Mesorhizobium* sp. Even after the decontaminating steps mentioned in the methods, these ASVs still represented the majority of the sequences for most larvae's communities. No relationship was found between DNA concentration and the abundance of these ASVs (Spearman's test, $r_s = 0.21$, $p = 0.34$ for ASV1 and $r_s = 0.09$,

p = 0.68 for ASV5). ASV1 was highly associated with sites T001 and T003 whereas ASV was mostly found in site T013 (Figure 2.6) even though samples from different sites were extracted and amplified together in a randomized design. In site T006, ASV1 was significantly more abundant in female samples than in male samples (*Mann–Whitney–Wilcoxon test*, U=*22*, p=*0.04).*

Figure 2.6: Mean relative abundance of ASV1 and ASV5 for spruce budworm bacterial communities at sampling sites.

2.5.2 Bacterial communities in gut tissues versus whole larvae

The bacterial communities of whole larvae were more diverse than those in gut tissues (Mann-Whitney-Wilcoxon, p=0.00416; Figure 2.7). In terms of community composition, different sample types clustered into two different groups (Figure 2.8). Variation in the composition of bacterial communities among whole larvae communities was greater than the variation among gut tissue samples, as shown by the larger dispersion of whole larvae communities in the NMDS ordination (Figure 2.8) and higher distance to centroid in general (Mann-Whitney-Wilcoxon on betadisper results, $p<0.001$; Figure 2.9). Community composition of gut tissue versus whole larval communities separated mainly on the first axis of variation, whereas variation among whole larvae samples was mainly along the second axis of variation. Type of sample, site and host-tree species explained 14%, 28% and 4% of the variation in community composition, respectively (PERMANOVA, p=0.001, p=0.001, p=0.016, respectively; Table 2.3), and the interaction between sample type and site explained 16% of the variation (p =0.001; Table 2.2). This means that the impact of site on bacterial community composition was partly dependent on whether it was part of whole larvae or gut tissue communities.

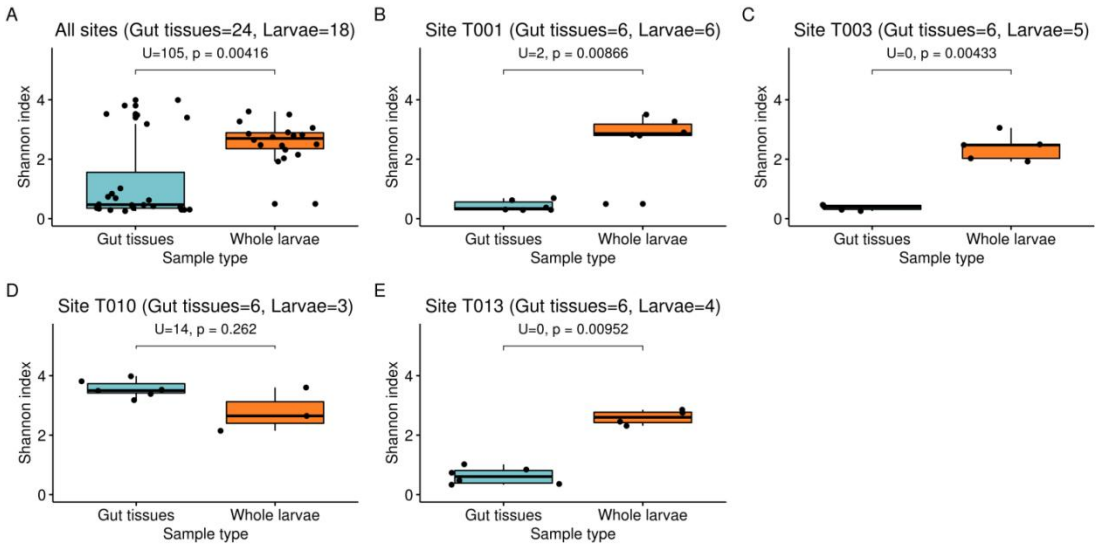

Figure 2.7: Boxplot showing the comparison between bacterial communities of spruce budworm whole larvae and gut tissue bacterial communities in terms of Shannon index. P-values of Mann-Whitney-Wilcoxon tests between larval and gut tissue community Shannon index are shown over each boxplot.

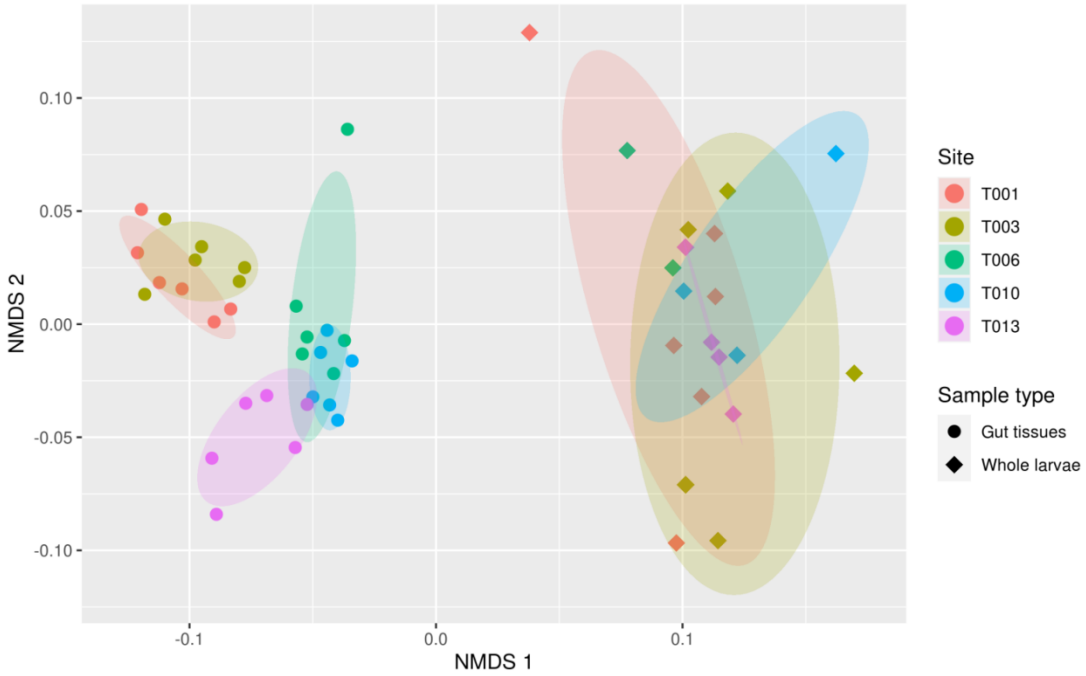

Figure 2.8: NMDS ordination of spruce budworm bacterial communities on Bray-Curtis dissimilarity for each site, whole larvae are shown as diamonds and gut tissue samples as circles. 68% confidence intervals ellipses representing sites are shown overlaid on the ordination.

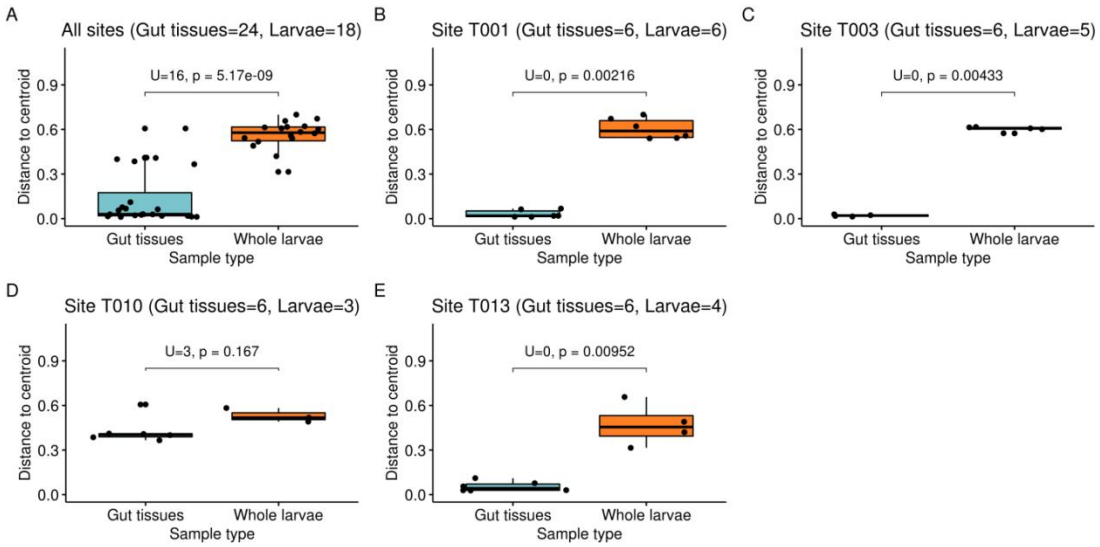

Figure 2.9: Boxplot showing the comparison between spruce budworm bacterial communities of whole larvae and gut tissues in terms of distance to group centroid in non-Euclidean distance measured by betadisper (Anderson et al., 2006). Groups were defined as combinations of sample type and site factors. P-values of Mann-Whitney-Wilcoxon tests between larval and gut tissue community distances to centroid are shown over each boxplot.

Table 2.3: Variation in spruce budworm bacterial community composition of gut tissue and larval communities as a function of sample type, sampling site and host-tree species, tested using a Bray-Curtis dissimilarity-based PERMANOVA.

	Df	Sum Sq	Mean Sq	F-Value	R2	P-Value
Sample type	1	2.78	2.78	13.97	0.141	0.001
Site	4	5.51	1.38	6.92	0.280	0.001
Host-tree species	2	0.76	0.38	1.92	0.039	0.016
Sample type*Site	3	3.11	1.04	5.20	0.158	0.001
Site*Host-tree species	7	1.76	0.25	1.26	0.089	0.094
Residuals	29	5.78	0.20		0.293	
Total	46	19.70			1	

When we separated samples of different types, we found that only microbe communities in gut tissues were correlated with site location, with site explaining the majority of variation in the composition of gut tissue communities (PERMANOVA*;* R^2=0.73, p =0.001; Table 2.4). At the phylum level, *Actinobacteriota* were much more abundant in gut tissue communities whereas *Acidobacteriota*, *Firmicutes* and *Proteobacteria* were more abundant in whole larvae (Figure 2.10). We identified 96 ASVs that were differentially abundant between gut tissue and whole larvae communities, with 25 of them being common to both datasets (Supp. Table 2.3). There were 59 ASVs positively associated with gut tissue communities belonging mostly in the phylum *Actinobacteriota*, whereas 37 ASVs were associated with whole larvae belonging mostly to the *Proteobacteria* phylum (Figure 2.10). The genera *Endobacter, Methylocella* and *1174-901-12* were all common to both datasets and highly associated with whole larvae. The highly abundant ASVs ASV1 and ASV5 mentioned earlier were absent in whole larvae, identified in Figure 2.10 as the two taxa with the lowest $\log_2$ fold change.

Table 2.4: Variation in spruce budworm bacterial community composition of gut tissue and larval samples as a function of sampling site and host-tree species nested within site, tested using a Bray-Curtis dissimilarity-based PERMANOVA, modeled separately for each sample type.

	Df	Sum Sq	Mean Sq	F-Value	R^2	P-Value
Whole larvae						
Site	3	1.19	0.40	1.00	0.170	0.445
Host-tree species within Site	4	1.82	0.45	1.15	0.261	0.139
Residuals	10	3.97	0.40		0.569	
Total	17	6.98			1	
Gut tissues						
Site	4	7.32	1.83	17.18	0.734	0.001
Host-tree species within Site	5	0.52	0.11	0.98	0.052	0.468
Residuals	20	2.13	0.11		0.214	
Total	29	9.98			1	

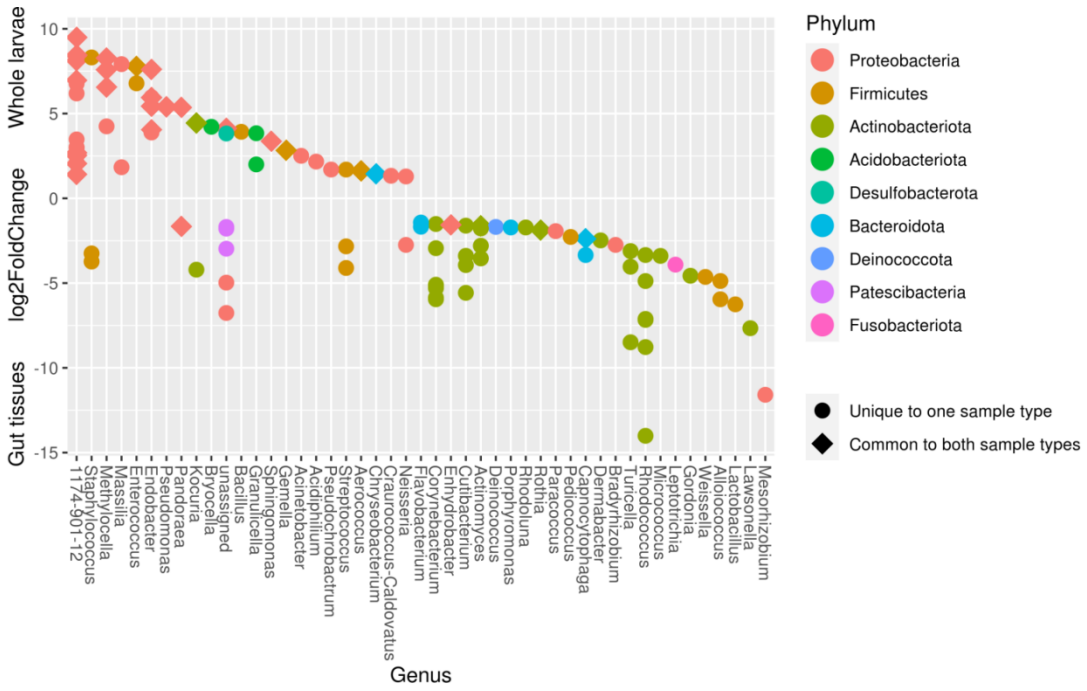

Figure 2.10: DeSEQ2 discriminant analysis of differentially abundant ASVs in spruce budworm bacterial communities of gut tissues and whole larvae. ASVs found in both larval and gut tissue communities are identified as diamonds whereas ASVs found in only one sample type are identified as circles.

2.6 Discussion

2.6.1 Larval sex differences

We did not detect an overall effect of sex on the composition of bacterial communities in the spruce budworm gut tissue, but there was a difference in the Shannon index of these communities between sexes. In terms of overall gut tissue community composition there were only a few weak and spatially inconsistent differences between the sexes. We did detect numerous bacterial taxa that were differentially abundant between male and female larvae but whose relative abundance was too low to drive significant variation in overall community composition. The

ASVs that differed in abundance between the sexes generally belonged to genera including *Corynebacterium*, *Staphylococcus, Gordonia*, *Actinomyces*, *Leptotrichia*, *Cutibacterium* and *Pediococcus*. Among these genera, *Pediococcus* is very abundant and potentially linked to sex in the mormon cricket *Anabrus simplex* (C. C. Smith et al., 2017). *Corynebacterium* is found in plant-eating Lepidoptera *Brithys crini*, and is part of the core microbiome of the apple buprestid *Agrilus mali* (Vilanova *et al.*, 2016, Zhang *et al.*, 2018). *Staphylococcus* and *Gordonia* were both found in the amaryllis borer *Brithys crini*'s gut tissues (Vilanova et al., 2016) whereas *Actinomyces* is associated with Hymenopteran gut tissues and with insecticide degradation in *Spodoptera frugiperda* (De Almeida et al., 2017; Matarrita-Carranza et al., 2017). *Cutibacterium* has been found in another insect pest, *Nilaparvata lugens*, and its abundance was associated with pest-vulnerable rice varieties compared to resistant varieties (Ojha & Zhang, 2019). *Leptotrichia*, to our knowledge, is the only one of these taxa which has never been associated with insect gut tissues, although the closely related bacterium *Sebaldella termitidis*, is a known resident of termites (Eisenberg et al., 2015). Most of these species were positively associated with females, suggesting that some microbiome related processes are specific to budworm females, a theory which deserves empirical testing. Even though overall gut tissue bacterial composition was not significantly related to larval sex, we identified sufficient differences in both bacterial diversity and abundance of specific taxa to lend support to our hypothesis that sex has an influence on spruce budworm gut tissue bacterial communities.

2.6.2 Bacterial communities in gut tissues versus whole larvae

The bacterial communities of spruce budworm gut tissues versus whole larvae were compositionally distinct. We found that site, and not host tree species, was the most important driver of microbial community structure in gut tissues. This contradicted our hypothesis that host tree species, which influences the diet and resulting gut tissue

microbiome structure (Landry et al., 2022), would be the most important driver of gut tissue microbiota composition; we expected that gut tissue bacteria would be more sensitive to the nature of the diet whereas the communities on whole larvae would be driven by the environment of the insect which is influenced by numerous spatial and environmental factors and varies among sites (Landry et al., 2022). Furthermore, the composition of microbial communities on the leaves that comprise the spruce budworm's diet has also been shown to vary both among host species and among sites (Laforest-Lapointe et al., 2016). This impact of sampling site on phyllosphere bacteria could explain the fact than gut tissue communities are very sensitive to sampling site, since any factor impacting microbes in their diet should directly impact what is found in their gut tissue. If we compare gut tissue communities with whole insects in terms of composition, it is clear that gut tissue communities are much more impacted by site location than whole larvae microbial communities. Whole larvae microbial communities were significantly more diverse and more variable within sites than gut tissue communities, as was expected since larvae are in contact with the environment and constantly exposed to potential colonizing microbes, whereas the lepidopteran gut tissue is a hostile environment acting as an ecological filter on potentially colonizing bacteria, and with dispersal limited to those bacteria that are ingested as part of feeding (Engel & Moran, 2013). Among bacteria common to both whole larvae and gut tissue communities, *Endobacter*, *Methylocella,* and *1174-901-12* were all positively associated with whole larvae and could have come from contact of the insect with foliage, where these taxa are known to be abundant (Khalifa et al., 2016; Kumar et al., 2019; Rousk & Bengtson, 2014).

At three out of five sites, bacterial gut tissue communities were completely dominated by one or two highly abundant ASVs that belong to the genera *Rhodococcus* and *Mesorhizobium*. The very clear clustering of gut tissues from these sites in terms of their community composition (Figure 2.8) seem to be strongly driven by the relative abundance of these two dominant ASVs. Both these taxa are known to be associated

with plants, and more specifically the rhizosphere of *Phragmites australis* (Pereira et al., 2015). The predominance of these two ASVs was unique to gut tissue samples suggesting that community assembly in the lepidopteran gut tissue selects for very specific taxa or leads to the possibility of competitive dominance by a small number of microbial taxa. In contrast, whole larvae communities were more diverse overall and less likely to be dominated by a handful of taxa, either due to differences in the selective environment or more likely due to the increased possibility of colonisation by diverse microbial taxa from the environment.

We are confident that the dominance of some samples by a small number of ASVs is not the result of contamination. First, all gut tissue dissections, DNA extractions, and PCR designs were randomized across sites and sampling dates so that samples from the same sites were not all processed or sequenced together. Second, negative controls from both the PCR and the DNA extraction were sequenced and analyzed using the *decontam* package (Davis et al., 2018) with the most stringent parameters to remove potential contaminant ASVs. As the dominance of these two ASVs persisted despite these precautions, it is probable that budworm gut tissue communities can be dominated by different taxa depending on spatial and environmental variation (captured by sampling sites in our study), whereas whole larval communities are both more diverse overall and less likely to be dominated by a few taxa. It is interesting that we observe a strong site effect, which was also found in another budworm microbiome study and could not be explained by spatial nor environmental factors (Landry et al. 2021). The lack of knowledge regarding what drives the important effect of site on budworm associated bacterial communities would need to be addressed to properly understand the relationship involving the budworm, bacteria, and their environment.

In terms of general composition, we found that whole larval and gut tissue communities of the spruce budworm are distinct and potentially responding to

different community assembly processes leading to differentially sensitivity to spatial and environmental factors. From a method perspective, this indicates that larval and gut tissue communities should not be used interchangeably when measuring insect microbiomes. In terms of biological explanations of these differences, we noted that inter-site variation was much higher in larval gut tissues, suggesting that gut tissue communities are potentially more sensitive to variation in the pool of colonizing microbes and that there is a greater possibility of competitive dominance of the gut tissue by a few microbial taxa, while the whole larval communities vary less spatially while being composed of a more diverse group of microbes, thus reflecting the importance of food resource and geographic variations in the phylosphere. Further studies are needed to compare the microbial communities of various anatomical parts to further elucidate how lepidopteran microbial communities are assembled in different parts of the insect body, and also to understand their impact on their host's biology, fitness, and population dynamics.

To compare communities of gut tissues and whole larvae, we used two datasets that came from samples collected in two different years that were amplified and sequenced separately, which could lead to a confounding of body part differences with temporal or sequencing batch effects (Pollock et al., 2018). However, we believe that there are several lines of evidence to support our interpretation that the differences we observed are due to true differences between communities of gut tissues versus whole larvae. To minimize confounding effects, samples from both groups were processed using the same protocols, by the same people, and their respective datasets were integrated in the very early stages of bioinformatics processing to minimize variation due to the difference between the two datasets. Also, studies of temporal variation in microbiomes in natural ecosystems have demonstrated very little inter-annual variation in microbiome structure (Cáliz et al., 2018; Epstein et al., 2019). Temporal variation has been shown to be relatively unimportant in studies of the foliage microbiome of the tree species that the spruce

budworm feeds on in Quebec (Laforest-Lapointe et al., 2016). It is therefore highly likely that the variation we observed between gut tissue and whole larvae communities is due to differences between these body parts, as has been demonstrated in other insects and mammals (Dekaboruah et al., 2020; Otti, 2015; Stehlikova et al., 2019).

2.7 Conclusions

Despite male and female budworm gut tissue communities being very similar overall in terms of composition, both Shannon index scores and the relative abundance of a few taxa showed measurable differences between sexes. Gut tissue communities varied strongly between collection sites, which indicates that the major drivers of budworm communities could be related to the insect's and the bacteria's environment. In contrast, there were strong and consistent differences in bacterial communities in spruce budworm gut tissues versus whole larvae samples, where gut tissue communities were much more variable among sites and were strongly dominated by two ASVs, whose prevalence was site dependent. This supports the hypothesis that an important part of spruce budworm bacterial communities lives outside of their gut tissues, and that the choice of body part to sample has important implications for the results of Lepidopteran bacterial communities.

2.8 Acknowledgements

We thank Jérémie Brisson and Marie-Line Buteau for contributing to sampling. We also thank Geneviève Bourret for assistance with sample processing and sequencing.

CHAPITRE III

VARIATION IN SPRUCE BUDWORM BACTERIAL ASSOCIATION NETWORKS AMONG WHOLE LARVAE AND LARVAL GUT TISSUES, HOST TREE SPECIES AND LARVAL SEX

Cet article est en préparation.

Mathieu Landry, premier auteur.
Département des sciences biologiques, Université du Québec à Montréal,
141 Av. du Président-Kennedy, Montréal, Québec, H2X 1Y4
landry.mathieu.5@courrier.uqam.ca

Patrick M.A. James
Institute of Forestry and Conservation, University of Toronto,
33 Willcocks St, Toronto, Ontario, M5S 3B2
patrick.james@utoronto.ca

Daniel Kneeshaw
Département des sciences biologiques, Université du Québec à Montréal,
141 Av. du Président-Kennedy, Montréal, Québec, H2X 1Y4
kneeshaw.daniel@uqam.ca

Steven W. Kembel
Département des sciences biologiques, Université du Québec à Montréal,
141 Av. du Président-Kennedy, Montréal, Québec, H2X 1Y4
kembel.steven_w@uqam.ca

3.1 Résumé

L'analyse des réseaux de co-occurrence microbienne permet de comprendre les facteurs déterminant les patrons complexes de co-occurrence et les interactions potentielles entre les taxons microbiens. Alors que les réseaux trophiques des animaux associés à l'important ravageur défoliateur de la tordeuse des bourgeons de l'épinette (TBE), *Choristoneura fumiferana*, ont déjà été caractérisés, les réseaux trophiques microbiens associés à cette espèce n'ont pas été étudiés, malgré leur importance potentielle pour la dynamique des épidémies de TBE. Des réseaux de co-occurrence bactérienne ont été inférés pour deux ensembles de données constitués de séquences de gènes d'ARNr 16S bactérien provenant de larves entières et de tissus intestinaux disséqués de TBE recueillis dans des zones d'éclosion au Québec et au Labrador. Nous avons caractérisé les réseaux d'associations bactériennes et leur variation en fonction du fait que les communautés provenaient de larves entières ou de tissus intestinaux, des espèces d'arbres hôtes et du sexe des larves. Nous avons découvert que les réseaux de co-occurrence bactérienne des tissus intestinaux des larves étaient organisés autour de *Rhodococcus* en tant que taxon moyeux, et que les communautés de tissus intestinaux étaient impliquées dans des réseaux plus denses que les réseaux de larves entières. Les réseaux de taxons bactériens provenant d'insectes récoltés sur le sapin baumier étaient en général organisés de manière plus centrale et plus dense que les réseaux de bactéries provenant d'insectes récoltés sur l'épinette noire. Les réseaux de différents sexes larvaires étaient similaires en termes de structure de réseau et de centralité, mais l'identité et l'importance de certains de leurs taxons moyeux respectifs variaient. Dans l'ensemble, il y avait une variation significative dans les réseaux de cooccurrence bactérienne de la TBE selon le type de communauté, les espèces d'arbres hôtes et le sexe des larves. Cet article représente la première caractérisation des réseaux de cooccurrence bactérienne associés à la TBE. Il consiste en une première étape pour comprendre comment les bactéries interagissent lorsqu'elles sont associées à un lépidoptère ravageur et de l'importance de différents facteurs pour expliquer la variation des réseaux de cooccurrence bactérienne.

3.2 Abstract

The analysis of microbial co-occurrence networks makes it possible to understand the drivers of complex patterns of co-occurrence and potential interactions among taxa. While the foodwebs of macroorganisms associated with the important defoliating pest spruce budworm (*Choristoneura fumiferana*) have been previously characterized, the structure of the microbial foodwebs associated with this species has not been studied, despite their potential importance for budworm outbreak dynamics. Bacterial co-occurrence networks were inferred for two datasets comprised of bacterial *16S* rRNA gene sequences from spruce budworm larvae and dissected gut tissues collected in outbreak areas in Quebec and Labrador. We analyzed sequence data to characterize microbial association networks and their variation as a function of whether communities were from whole larvae or larval gut tissues, host tree species, and larval sex. We found that networks of bacterial co-occurrence of larval gut tissues were organized around *Rhodococcus* as a hub taxon, and that gut tissue communities were involved in networks denser than networks of whole larvae. Networks of bacterial taxa from insects collected on balsam fir were in general more centrally organized and denser than networks of bacteria from insects collected on black spruce. Networks from different larval sexes were similar in terms of network structure and centrality, but the identity and importance of some of their respective hub taxa varied. Overall, there was significant variation in budworm bacterial co-occurrence networks along community type, host tree species, and larval sex.This article represents the first characterization of bacterial co-occurrence networks associated with the spruce budworm, providing a first step towards understanding how bacteria interact together when in association with a lepidopteran pest and the importance of different factors to explain variation in bacterial co-occurrence networks.

3.3 Introduction

Network analysis provides useful tools to visualize and analyze interactions within a community, whether they are social media, food webs or microbiomes (Dunne et al., 2002; Freeman, 1978; Kishore et al., 2020). For network analysis in the context of microbial ecology, microbial taxa are represented as nodes and co-occurrence between them as edges (Albert & Barabási, 2002). Co-occurrence networks can capture the complexity of microbial communities and their interactions in a way that compositional analysis which consists of comparing relative abundances between locations, cannot (Layeghifard et al., 2017). Taxa with a particularly high number of interactions within their networks, often called hub taxa, are thought to be of great importance to the structure and function of the community and the host (Agler et al., 2016), although there is debate about how to accurately identify hub taxa using network analysis (Delmas et al., 2019). Microbes are known to cooperate with closely and distantly related organisms, forming emergent assemblages (e.g., biofilms) in which microbes function as more than the sum of their parts. These instances of cooperation can be missed by looking only at the relative abundance of taxa in a community rather than the full network of positive and negative co-occurrences (Dalton et al., 2011; Murray et al., 2014).

Network analysis can potentially give insights into the mechanisms that explain relative abundance patterns, whether they involve abiotic factors, the microbes themselves, or the host (Agler et al., 2016). Because observing direct interactions between microbes at a community scale is impossible using commonly-employed methods such as sequencing of barcode genes, microbial interaction networks are usually inferred from co-occurrence analysis (van der Heijden & Hartmann, 2016). Co-occurrence network analysis measures positive and negative associations between the relative abundances of pairs of taxa at specific locations and infers interactions from these co-occurrence networks (Friedman & Alm, 2012). However, interactions

cannot always be inferred from co-occurrence (Blanchet et al., 2020). With this in mind, co-occurrence network analysis serves to generate hypotheses that can be tested using mechanistic experiments; for example, positive co-occurrence could indicate the presence of mutualistic interactions, whereas negative co-occurrence could indicate antagonistic interactions (Delmas et al., 2019).

Multiple metrics derived from network science can be used to measure the characteristics of microbial networks or of individual nodes within the network. Commonly used network metrics include the total number of nodes and edges, density, diameter and motifs numbers. Number of nodes, also referred to as order, indicates the number of taxa in a community for which their abundance was correlated with the abundance of at least another taxon. Number of edges, also known as size, indicates the total number of co-occurrence associations in the network, and can be divided by the order to calculate the linkage density, i.e., the average number of edges each taxon possesses in the network. Diameter is the length of the shortest path between the two most distant nodes in a network. Network diameter has been hypothesized to be representative of the sensitivity of a community to a perturbation (Delmas et al., 2019). Motifs are recurrent well-connected subgraphs involving 3 or 4 nodes representing patterns of co-occurrence and their over representation can be indicative of closer microbial cooperation patterns (Milo et al., 2002; Simmons et al., 2019).

In addition to whole-network metrics, several network metrics for individual nodes have been proposed such as node centrality. Centrality measures are useful to identify well-connected or influential nodes, called hub taxa, which could be of particular importance to the network and by association the community and the host (Agler et al., 2016; Delmas et al., 2019). Commonly used centrality measures include (1) betweenness centrality, which evaluates how often a node is a bridge between two others and (2) eigenvector centrality, which evaluates node influence by scaling

scores depending on the score of influenced nodes (Bonacich, 1987; Delmas et al., 2019; Freeman, 1977, 1978).

Environmental microbiome network analysis is still recent and the knowledge we possess about microbial networks in nature is relatively scarce (Delmas et al., 2019; Layeghifard et al., 2017). Legrand et al. (2019) found that soils whose microbial networks were more densely connected were more resistant to a fungus pathogen, *Fusarium graminearum*. Soil bacterial networks were also found to be less compartmentalized in high stress environments and the more stressful the environment, the more positive co-occurrence dominated over negative co-occurrence (Hernandez et al., 2021). In American forests, soil microbes form networks that are sensitive to biogeographical gradients, with network metrics varying along environmental and spatial gradients (Tu et al., 2020). Bacterial networks in sand biofilters showed that taxa with low relative abundance could still be very well-connected hub taxa within a network (Zhao et al., 2021). In *Arabidopsis thaliana*-associated microbial networks, multiple microbes were identified as hub taxa and it was hypothesized that host effects were propagated through these hub taxa to the rest of the community (Agler et al., 2016). In insects, the apple tree borer *Agrilus mali* hosts a microbial network in which *Kocuria* consistently co-occurred with *Corynebacterium* and in which multiple complex co-occurrence pattern were found between bacteria and fungi (Zhang et al., 2018). Insofar as insects have fundamental impacts on human and ecological systems (Dimopoulos, 2003; Hung et al., 2018; Wright, 1986), characterizing microbial co-occurrence networks hosted within insects could help explain the significant level of variability in their microbial communities (Engel et al., 2012; Engel & Moran, 2013), and eventually lead to a more general understanding of the factors governing insect microbiome composition.

Lepidoptera, which include some of the most damaging defoliating insects (Pantoja et al., 1986; Piper et al., 2015; D. A. Russell et al., 1993), host microbial communities

for which the determining factors are not generally well-known. The existence of a resident microbiome in Lepidoptera has been called into question, potentially meaning that diet determined the composition of bacteria present in these insects (Hammer et al., 2017). The eastern spruce budworm, *Choristoneura fumiferana*, is a major defoliating pest that attacks multiple conifers species, among them balsam fir (*Abies balsamea*) and black spruce (*Picea mariana*) over northeastern America (Royama, 1984). Spruce budworm outbreak periodicity is around 30 years. During outbreaks, feeding larvae cause tremendous damage to fir and spruce trees. Almost 13,500,000 hectares for northern forest has been defoliated in the current outbreak (Ministère des Forêts, 2020). We do not fully understand the factors controlling outbreaks, nor exactly how the main biological agent to control the budworm, *Bacillus thuringiensis* toxins, kill the insect when ingested (Frankenhuyzen et al., 2010; Pureswaran et al., 2016). The budworm is part of a complex food web involving parasitoid insects and avian predators that influence its population and outbreak dynamics (Crawford & Jennings, 1989; Eveleigh et al., 2007; Royama, 1984). Bacterial communities could be an important uncharacterized part of the budworm's foodweb, as they are for other insect systems (Breznak, 1982; Calderón-Cortés et al., 2011; A. H. Smith et al., 2015; Warnecke et al., 2007).

Budworm whole larvae microbiome composition depends on site and host-tree species, with microbial communities of foliage eaten by the budworm closely mirroring these variations (Landry et al. Chapter 1). Bacterial communities found in the budworm's digestive track are significantly differently from communities hosted in whole larvae from the same locations, in terms of diversity, composition and spatial variability (Landry et al. Chapter 2). We know that budworm gut bacterial communities change depending on rearing conditions (Landry et al., 2015) and that there are subtle differences between the gut tissue communities of different sexes (Landry et al. Chapter 2). However, to date the network structure of the budworm microbiome has not been quantified. Characterizing the microbial co-occurrence

networks behind the patterns observed in budworm microbiomes could reveal important bacterial associations and hub taxa driving community variation and potentially affecting the budworm and the distribution of its outbreaks.

In this study, we quantified the ecological network structure of spruce budworm whole larval and gut tissue bacterial communities to test hypotheses about the influence of physiological structure, host-tree species, and larval sex for driving the structure and hub taxa of ecological association networks.

Previous studies have quantified insect microbiomes associated with either the whole insect gut tissue tissues only (Engel & Moran, 2013; Hammer et al., 2014) but how microbial associations vary between gut tissues and whole larvae is still unknown. We test three hypotheses related to differences between these sample sources: 1) Because the gut tissue is physically nested within the larval body, we predict that microbial co-occurrences within gut tissue communities will be nested within those of whole larvae networks. 2) Because whole larvae are more exposed to the environment relative to gut tissues, we predict whole larvae microbial networks will include more diverse taxa resulting in higher network order and lower network diameter. 3) Because two taxa (*Rhodococcus* and *Mesorhizobium*) dominated gut tissue communities in terms of relative abundance, we predict that these species will have the highest measure of network centrality in gut tissue co-occurrence networks (Landry et al. Chapter 2).

Balsam fir and black spruce are among the budworm's main host tree species (Royama, 1984) and important differences have been found in the communities of insect feeding on different host tree species (Landry et al., 2015; Landry et al., 2021). Here, we test the hypothesis that: 4) microbial networks differ between whole larvae and gut tissue communities within different host species. Although one might expect diet (i.e., foliage) to affect gut tissue communities more than whole larvae

communities, the opposite trend was observed (Landry et al. Chapter 2) and so we predict differences in network structure and metrics to be more marked in whole larvae because of the selective nature of the budworm gut tissue (Engel & Moran, 2013).

Communities of gut tissues from larvae of different sexes were very similar overall in the relative abundance of most taxa, with a few relatively rare specific taxa being differentially abundant between the sexes (Landry et al. Chapter 2). 5) We consequently predicted no difference in hub taxa between sexes.

3.4 Methods

To address our questions about the roles of community type, host tree species, and larval sex on microbial co-occurrence networks, we analysed data from published data sets on eastern spruce budworm larval gut tissue (Landry et al. Chapter 2) and whole larval (Landry et al. Chapter 1) bacterial community composition. The effect of host tree species was measured separately in whole larvae and gut tissue networks.

3.4.1 Whole larvae dataset

The sampling sites for whole larvae comprised of 26 sites over North Coast of Quebec, Gaspesie and Labrador (Supp. Figure 3.1). In each site were sampled six randomly selected trees (three black spruce and three balsam fir) for three larval samples per tree that were collected from each tree using aseptic technique and stored in sterile containers. Samples were stored in RNAlater on ice for approximately 2 days, then shipped to Montreal where they were stored at -80^{o}C prior to processing and sequencing.

3.4.2 Gut tissues dataset

Larvae used for the characterization of gut tissue communities were collected at five sampling sites in Côte-Nord, Quebec (Supp. Figure 3.2). As for whole larvae, six randomly selected trees (three black spruce and three balsam fir) were sampled at each site for three larval samples per tree that were collected from each tree using aseptic technique. Five sites (T001, T003, T006, T010 and T013) overlapped over the two datasets. Samples were stored and shipped in the same way as whole larvae. Larvae were dissected aseptically to identify larval sex and sample the gut tissue.

3.4.3 Sequencing and sequence data processing

Samples were processed and sequences were generated according to Landry et al. (Chapter 2). Briefly, larvae and dissected larval gut tissues were processed using the Qiagen PowerSoil DNA extraction kit after which we amplified the *16S* rRNA gene variable regions V5-V6 (799F-1193R primers) (Chelius & Triplett, 2001) and added barcodes and Illumina adapters for sequencing with Illumina MiSeq V3 technology (paired end 2x300 nucleotide reads). DNA extraction and PCR were performed in a randomized design where samples were processed independently of their site, host-tree and sex for gut tissue and whole larvae samples separately and extraction (water processed with the extraction kit) and PCR negatives were incorporated in the analysis. We curated, merged, and clustered sequences into amplicon sequence variants (ASVs) using default parameters, removed chimeras according to DADA2 1.12.1 package (Callahan et al., 2016) standard operating procedures with the “consensus” method, and assigned taxonomy to ASVs by comparison with the SILVA database version 128 (Quast et al., 2012) in R software 3.6.0(R Development Core Team, 2013). The whole larvae dataset contained of 5,649,507 reads over 3619 ASVs in 258 samples, for an average of 21,897 reads per sample whereas the gut tissues dataset was made of 2,866,905 reads over 1308 ASVs in 119 samples, for an average of 24,092 reads per sample.

3.4.4 Removal of contaminant sequences

In order to remove potential contaminant ASVs that originate from contamination during the sequencing process as opposed to ASVs that were truly present in the environmental samples, we used a contaminant ASV removal approach. 11 extraction and 11 PCR negative control samples were included in the whole larvae dataset whereas the gut tissues dataset contained 3 extraction and 2 PCR negative control samples. The *decontam* 1.1.2 package (Davis et al., 2018) was used for identification of ASVs associated with negative control samples as contaminants. ASVs associated with negative DNA extraction and PCR controls were removed from datasets in function of their respective extraction and PCR controls. Contaminant ASVs detection was performed according to the prevalence method with an aggressive prevalence threshold (0.5) to ensure the ASVs who appeared in more controls than true samples were identified as contaminants. Overall, in the whole larvae dataset, 483 ASVs representing 1,778,258 reads were removed from the analysis and 107,755 reads belonging in 97 ASVs were removed in the gut tissues dataset.

3.4.5 Network inference

ASVs were aggregated to the taxonomic genus level to reduce network complexity and limit variability due to rarefaction. Each experimental group had its own community dataset, which was rarefied to 3000 reads per sample, for a total of 360,000 reads in 120 whole larvae samples and 270,000 reads in 90 gut tissue samples. We constructed rarefaction curves to determine if the sequencing depth (and sampling effort) was sufficient to sample a significant proportion of the microbial diversity both in spruce budworm gut tissues and whole spruce budworm samples. For each experimental group, we filtered out genera that occurred in fewer than five samples belonging to that group. We used the *sparcc* suite from the package SpiecEasi version 1.1.0 for subsequent co-occurrence calculation and network inference made especially to account for sparsity bias in compositional data

(Friedman & Alm, 2012; Kurtz et al., 2015). Co-occurrence between every pair of genera was calculated with *sparcc* with default parameters by estimating Pearson correlations between log-transformed abundances. Community tables were randomized by permutation and *p*-values were calculated for each pair of genera with a setting of 1000 bootstraps to measure whether co-occurrence values observed differed from co-occurrence in randomly generated communities. Pairs of genera whose *p*-values, after correction with the Benjamini Hochberg method (Benjamini & Hochberg, 1995), were under 0.05 were kept in an adjacency matrix that was converted to a network object with default parameters for subsequent network manipulation using the igraph package 1.2.6 (Csardi & Nepusz, 2006) in R. For each network, we generated 999 null models by randomly permuting edges while keeping the total number of edges constant and recalculating all network and node metrics to evaluate how network properties differed from random co-occurrence patterns (Gotelli & Graves, 1996; Gotelli, 2001). Network and node metric distributions from null models were compared with metrics from our inferred networks to assess how the characteristics of each network metric differed from the random expectation. Nodes were identified as hub taxa when either their betweenness centrality or eigenvector centrality score was higher than expected by chance (outside 95% confidence interval).

3.5 Results

3.5.1 Whole larvae and larval gut tissues

Both whole larvae and larval gut tissue networks showed organization around some central genera dominated by *Actinobacteriota* and *Proteobacteria*, with a few peripheral networks (Figure 3.1). Whole larvae networks involved much more bacterial genera (70 to gut tissue network's 40) for a similar number of edges, resulting in gut tissue networks being much denser (3.73 Linkage density compared

to 1.86, Table 3.1). Only one edge was found to be common to both whole larvae and gut tissue networks, which was the negative co-occurrence between *Acidiphilium* and *Lawsonella*. When taking into account both centralized measures of centrality (*betweenness* and *eigenvector*), the whole larvae network was more central than the gut tissues network, with both networks being more central than expected by chance (outside 95% confidence interval) for both measures (Figure 3.2). Both networks had a higher number of motifs than expected by chance (outside 95% confidence interval), with the gut tissues network involving a markedly higher number of motifs (Figure 3.2). Diameter was remarkably lower in the gut tissues network that in the whole larvae network but higher than expected by chance (outside 95% confidence interval), contrary to the whole larvae network. Both networks had very few hub taxa in terms of eigenvector centrality, with only an unidentified genus of *Caulobacteraceae* and *Rhodococcus* being identified as hub taxa (outside 95% confidence interval), for whole larvae and gut tissues respectively (Figure 3.3). In terms of betweenness centrality, hub taxa (outside 95% confidence interval) were *Bacillus*, *Terriglobus* and an unassigned genus of *Acetobacteraceae* in the whole larvae network and were *Rhodococcus*, *Bradyrhizobium* and *Streptococcus* in the gut tissues network (Figure 3.3).

Figure 3.1: Co-occurrence network of bacterial communities in spruce budworm larvae and gut tissues collected from Quebec and Labrador and Newfoundland. Co-occurrence between every pair of genera was calculated with sparcc with default parameters by estimating Pearson correlations between log-transformed abundances. Community tables were randomized by permutation and p-values were calculated for each pair of genera with a setting of 1000 bootstraps to measure whether co-occurrence values observed differed from co-occurrence in randomly generated communities. Full edges indicate positive significant ($p < 0.05$) co-occurrences between genera. Dotted edges indicate negative significant ($p < 0.05$) co-occurrences between genera. Node sizes are proportional to relative abundance. Node colors represent phylum-level taxonomy. Genera whose co-occurrence is present in networks of both groups are labeled.

Table 3.1: Proportion of positive and negative co-occurrences, network size, order and linkage density for networks of bacterial taxa collected from spruce budworm in Quebec. Co-occurrence networks were based on co-occurrence between every pair of genera calculated with sparcc by estimating Pearson correlations between log-transformed abundances.

Network	Positive co-occurrences	Negative co-occurrences	% of positive co-occurrences	Size	Order	Linkage density
Gut tissues	62	87	41.61	149	40	3.73
Whole larvae	56	74	43.08	130	70	1.86
Gut tissues/Balsam fir	31	41	43.06	72	29	2.48
Gut tissues/Black spruce	19	26	42.22	45	23	1.96
Whole larvae/Balsam fir	22	33	40	55	41	1.34
Whole larvae/Black Spruce	21	30	41.18	51	43	1.19
Male	35	34	50.72	69	32	2.16
Female	26	37	41.27	63	28	2.25

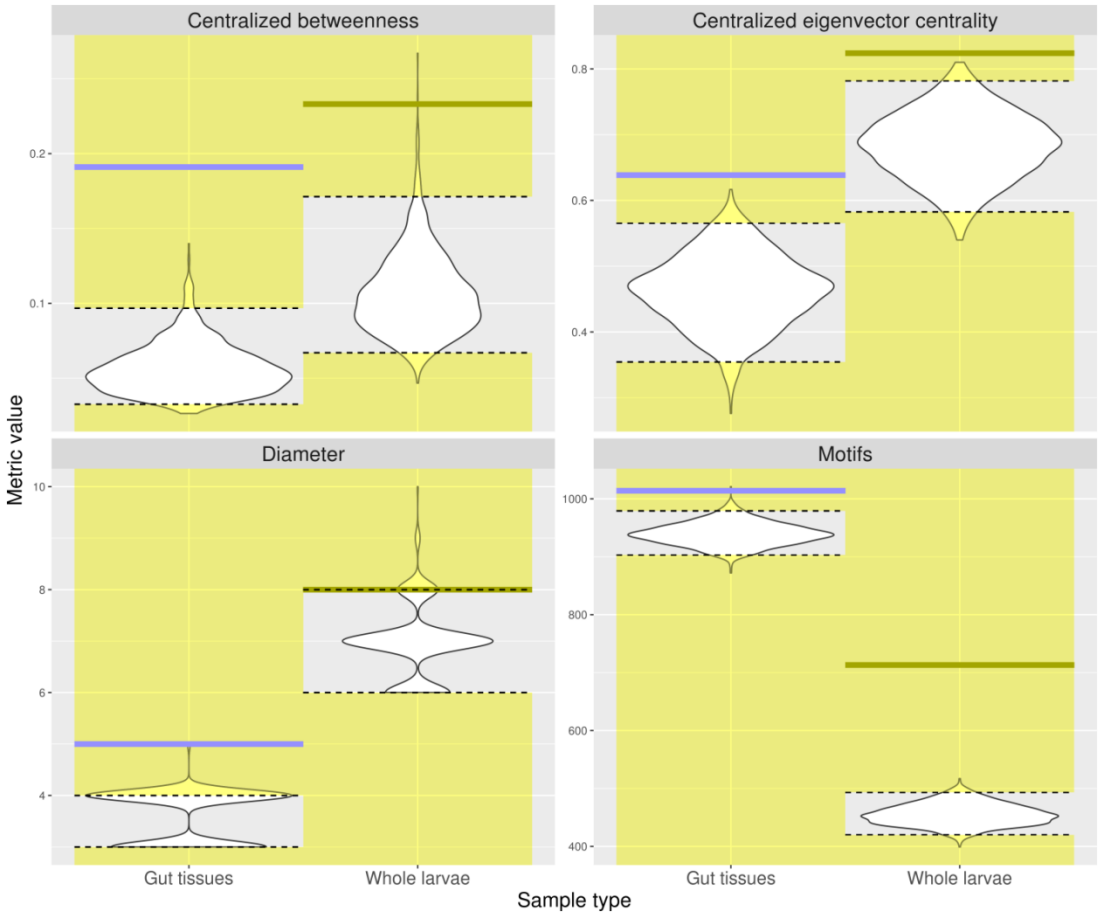

Figure 3.2: Proportion of positive and negative co-occurrences, network size, order and linkage density for networks of bacterial taxa collected from spruce budworm in Quebec. Co-occurrence networks were based on co-occurrence between every pair of genera calculated with sparcc by estimating Pearson correlations between log-transformed abundances.

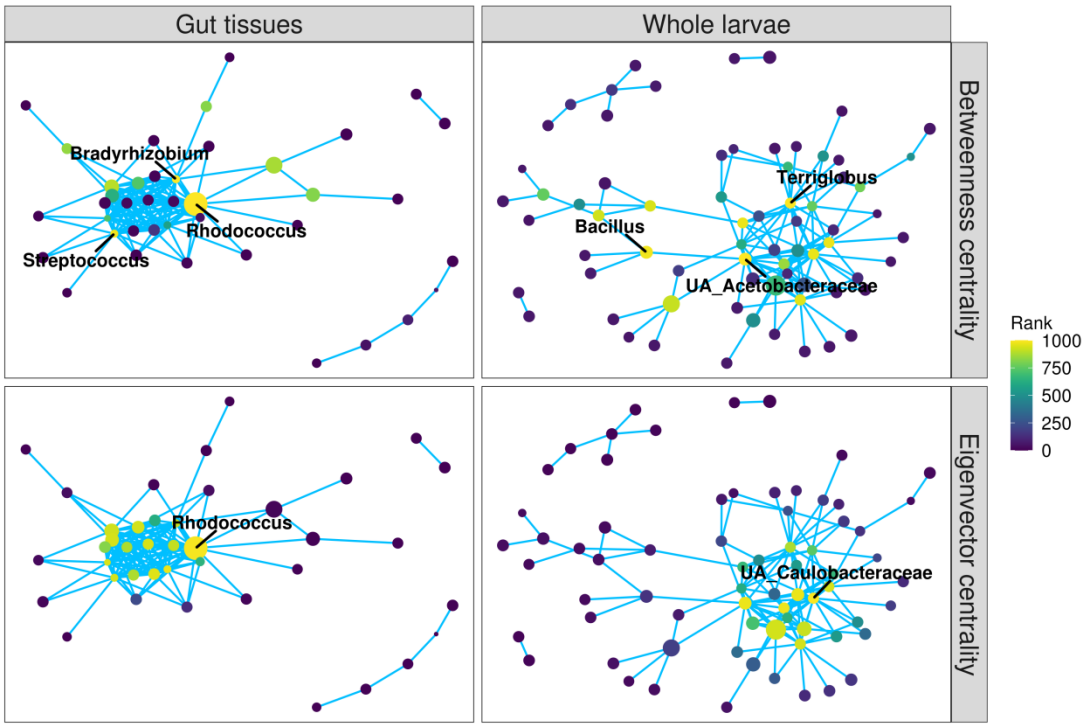

Figure 3.3: Graph representation of each genus for each network in networks of spruce budworm larvae and gut tissues, where vertices are shaded depending on their ranks in the null model distribution and their size is proportional to relative abundance. Only genera outside the 95% confidence interval are labeled.

3.5.2 Host tree species variation

With both gut tissue and whole larvae, networks associated with balsam fir as a host-tree species were denser than those associated with black spruce (Figure 3.4, Table 3.1). The number of motifs and centralized betweenness centrality were consistently lower in black spruce networks (Figure 3.5). Except for diameter, all network metrics varied similarly between host tree species independently of whether networks were inferred from whole larvae or gut tissue communities (Figure 3.5). In whole larvae networks, the diameter of the black spruce network was almost double that of the balsam fir network whereas in gut tissue networks diameter did not vary between host tree species (Figure 3.5). In the case of node metrics, host tree species had a small impact on hub taxa (Figure 3.6). Gut tissue networks had *Rhodococcus* and *Lawsonella* as hub taxa on balsam fir whereas the black spruce network had only *Corynebacterium* (Figure 3.6). In whole larvae networks, hub taxa went from unassigned *Caulobacteraceae* in the balsam fir network to both *Caulobacteraceae* and *Acetobacteraceae* in the black spruce network (Figure 3.6).

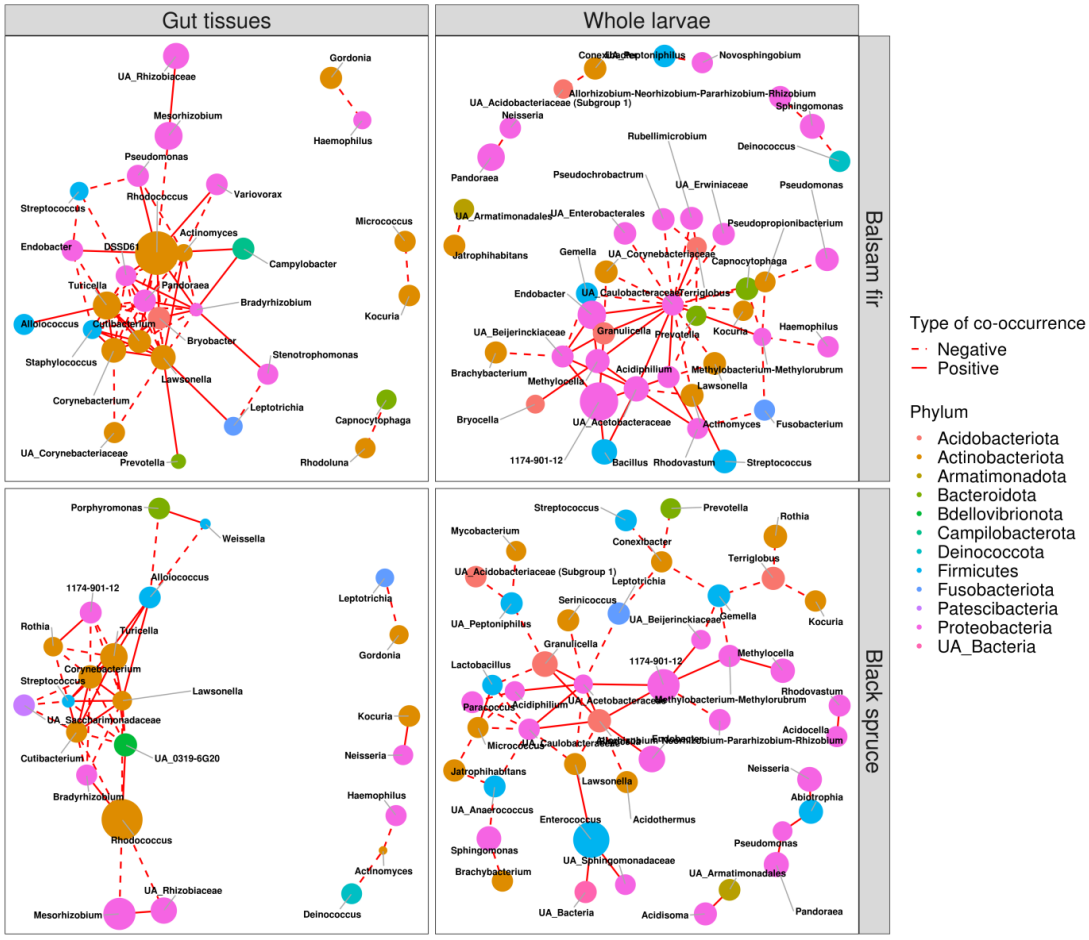

Figure 3.4: Co-occurrence network of bacterial communities in spruce budworm larvae and gut tissues feeding on two different host-tree species collected from Quebec and Labrador and Newfoundland. Co-occurrence between every pair of genera was calculated with sparcc with default parameters by estimating Pearson correlations between log-transformed abundances. Community tables were randomized by permutation and p-values were calculated for each pair of genera with a setting of 1000 bootstraps to measure whether co-occurrence values observed differed from co-occurrence in randomly generated communities. Full edges indicate positive significant ($p < 0.05$) co-occurrences between genera. Dotted edges indicate negative significant ($p < 0.05$) co-occurrences between genera. Node sizes are proportional to relative abundance. Node colors represent phylum-level taxonomy.

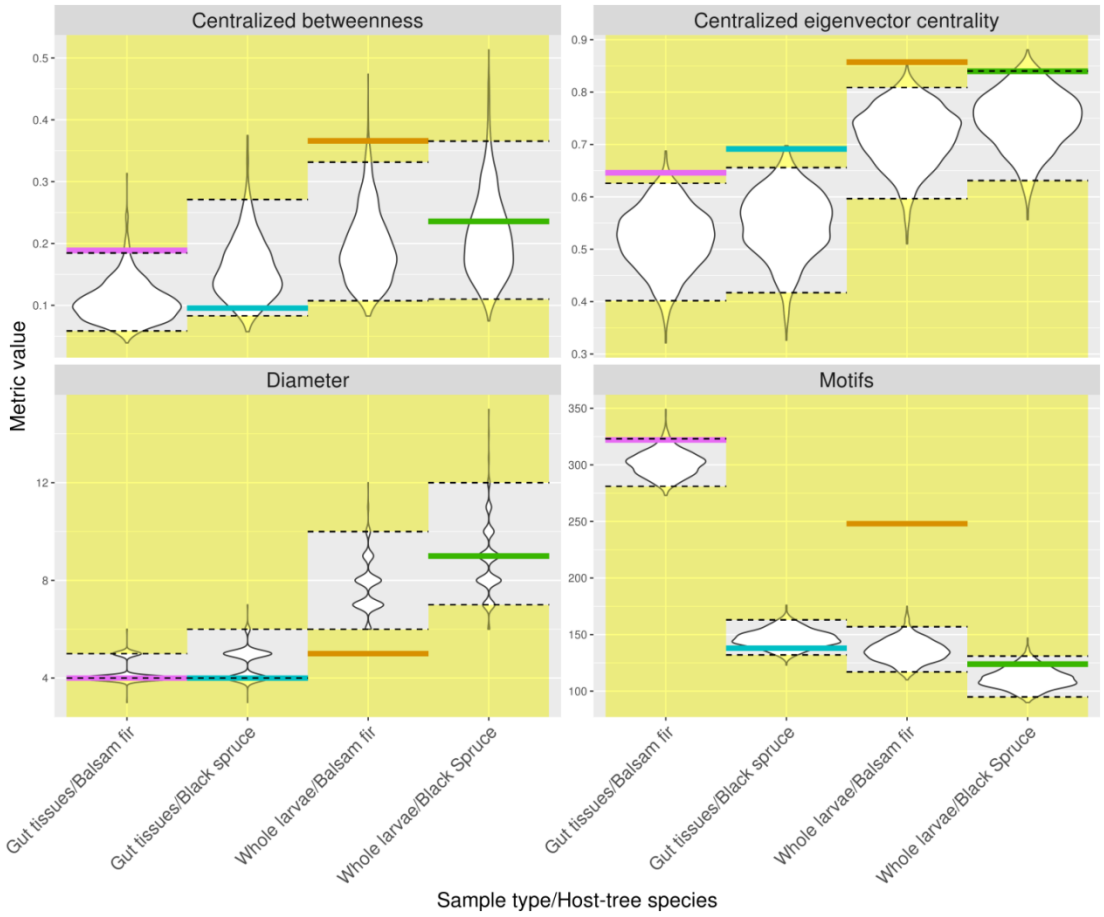

Figure 3.5: Violin plot of measured network metrics compared to null model distribution in networks of spruce budworm larvae and gut tissues feeding on two different host-tree species. The white shapes are the null distributions, the colored lines are the observed values for our networks (purple for the balsam fir gut tissue network, cyan for the black spruce gut tissue network, orange for the balsam fir whole larvae network and green for the black spruce whole larvae network). Yellow shading areas represent values outside 95% confidence intervals (with 999 null models, ranks either under 25 or over 975 out of 1000).

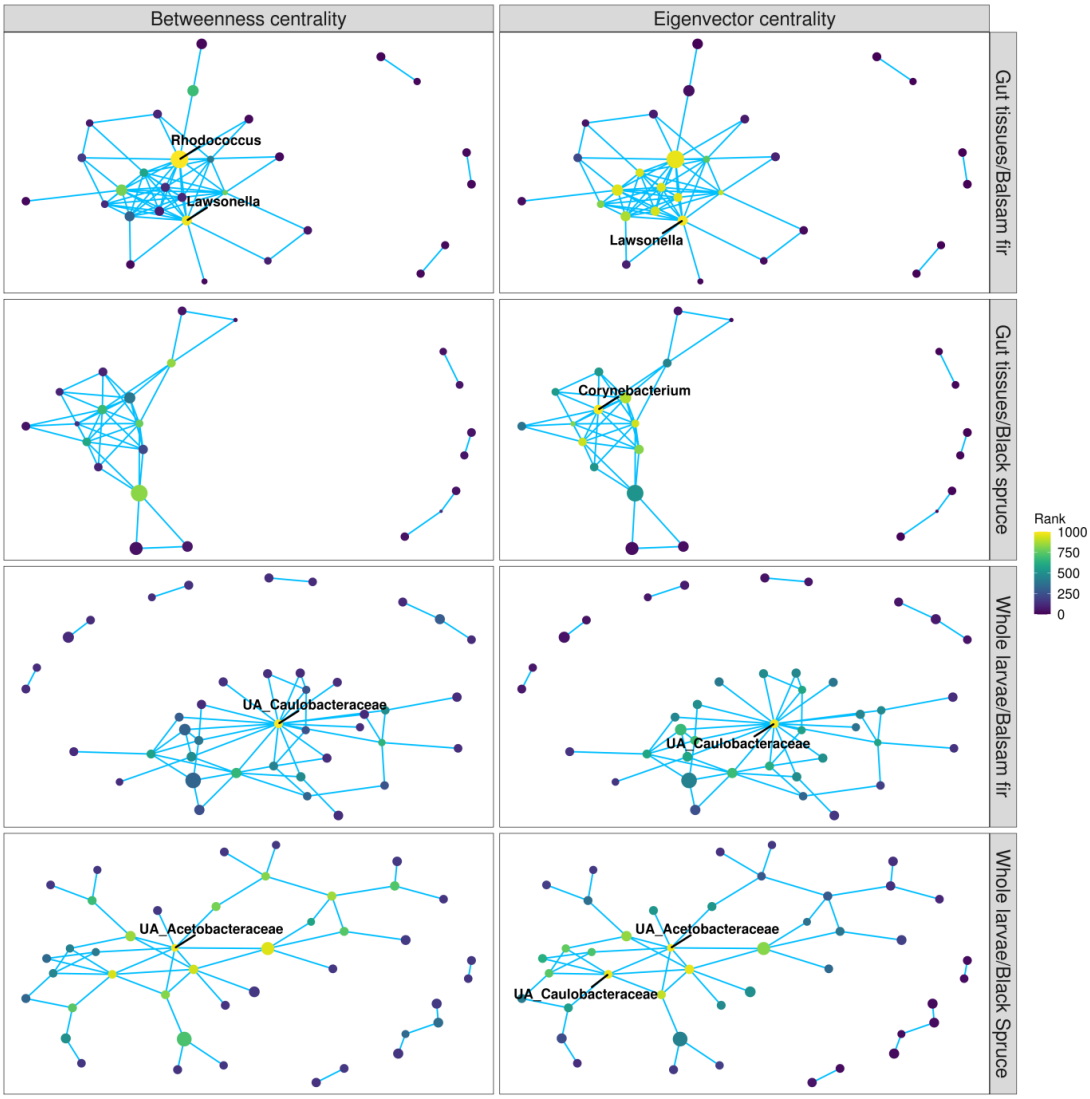

Figure 3.6: Graph representation of each genus for each network in networks of spruce budworm larvae and gut tissues feeding on two different host-tree species, where vertices are shaded depending on their ranks in the null model distribution and their size is proportional to relative abundance. Only genera outside the 95% confidence interval are labeled.

3.5.3 Larval sex variation

Networks of communities of males and females were very similar, with both being dominated by a group centered around *Rhodococcus* and other *Actinobacteriota* (Figure 3.7) and both having networks of similar density (Table 3.1). Of all networks presented in this study, networks of male budworm gut tissues had this highest percentage of co-occurrences being positive, reaching 50.7% compared to the other networks averaging 41.7% positive co-occcurrences. Network metrics were also similar between sexes for all metrics except centralized betweenness centrality. Both networks also showed higher centralized betweenness centrality, centralized eigenvector centrality and diameter than expected by chance (outside 95% confidence interval) (Figure 3.8). In the case of motifs, the network of female gut tissues was both higher than the male gut tissues network and that expected by chance (outside 95% confidence interval, Figure 3.8). *Lawsonella* and *Streptococcus* were hub taxa exclusively in networks of female budworm whereas *Turicella*, *Enterococcus*, *Rhizobiaceae* and *Capnocytophaga* were hub taxa only in networks of male gut tissues (Figure 3.9).

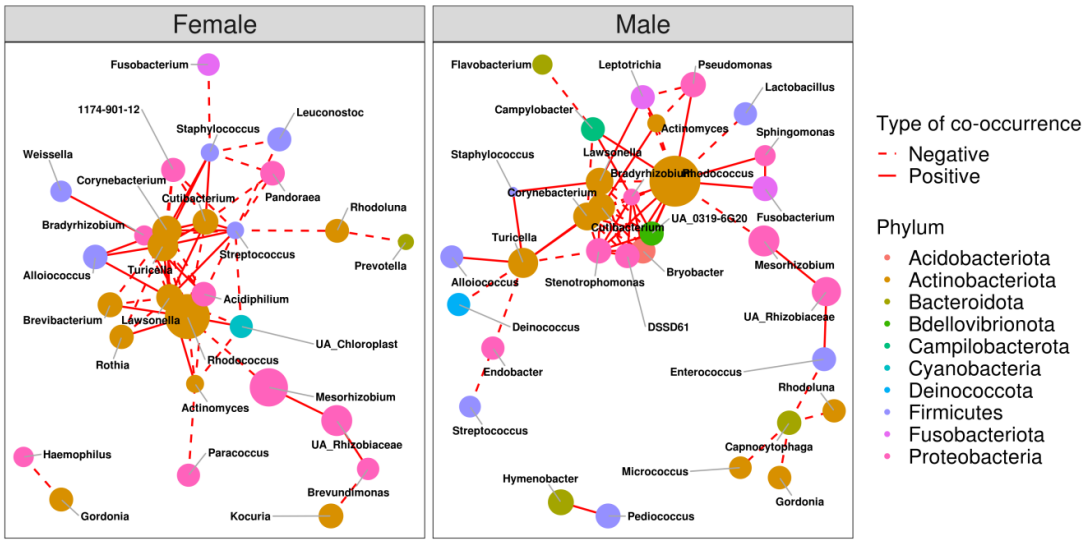

Figure 3.7: Co-occurrence network of bacterial communities in spruce budworm gut tissues from larvae of different sexes collected from Quebec. Co-occurrence between every pair of genera was calculated with sparcc with default parameters by estimating Pearson correlations between log-transformed abundances. Community tables were randomized by permutation and p-values were calculated for each pair of genera with a setting of 1000 bootstraps to measure whether co-occurrence values observed differed from co-occurrence in randomly generated communities. Full edges indicate positive significant ($p < 0.05$) co-occurrences between genera. Dotted edges indicate negative significant ($p < 0.05$) co-occurrences between genera. Node sizes are proportional to relative abundance. Node colors represent phylum-level taxonomy.

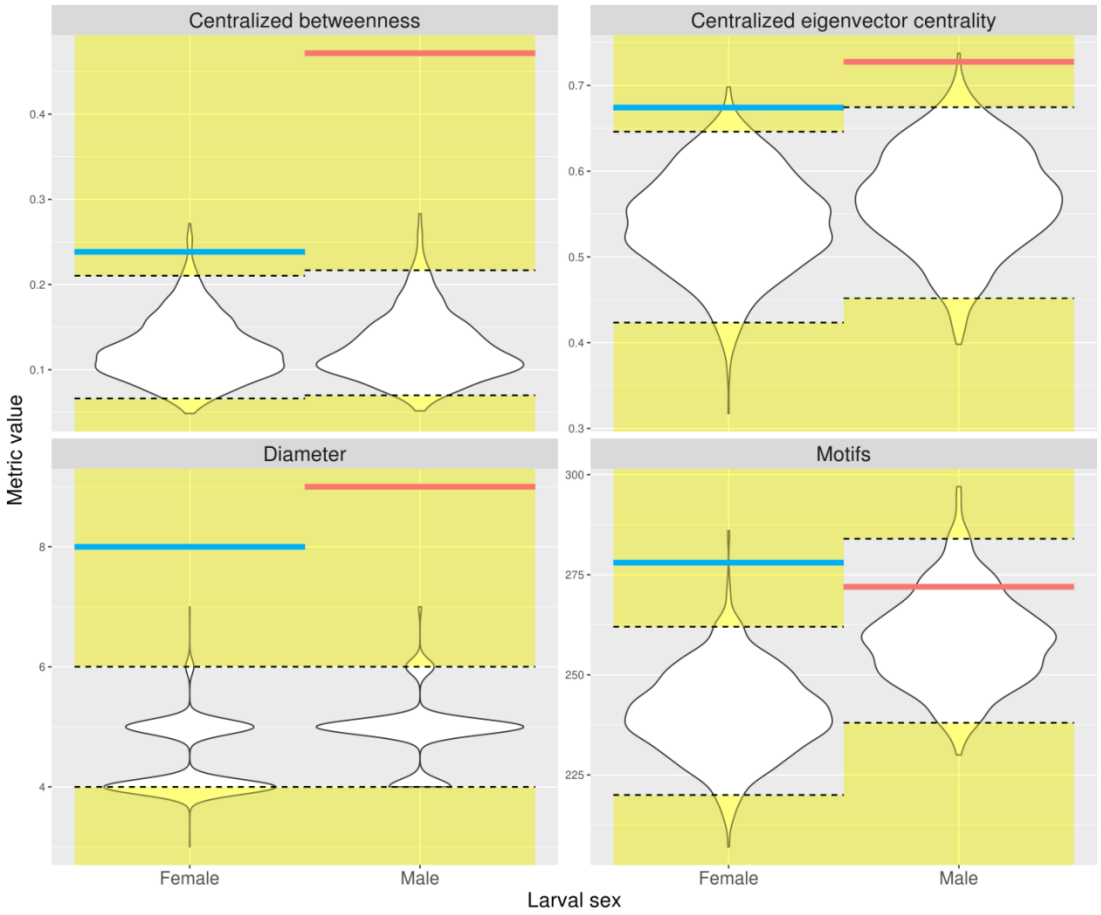

Figure 3.8: Violin plot of measured network metrics compared to null model distribution in networks of spruce budworm gut tissues from larvae of different sexes. The white shapes are the null distributions, the colored lines are the observed values for our networks (light blue for the female gut tissue network and red for the male gut tissue network). Yellow shading areas represent values outside 95% confidence intervals (with 999 null models, ranks either under 25 or over 975 out of 1000).

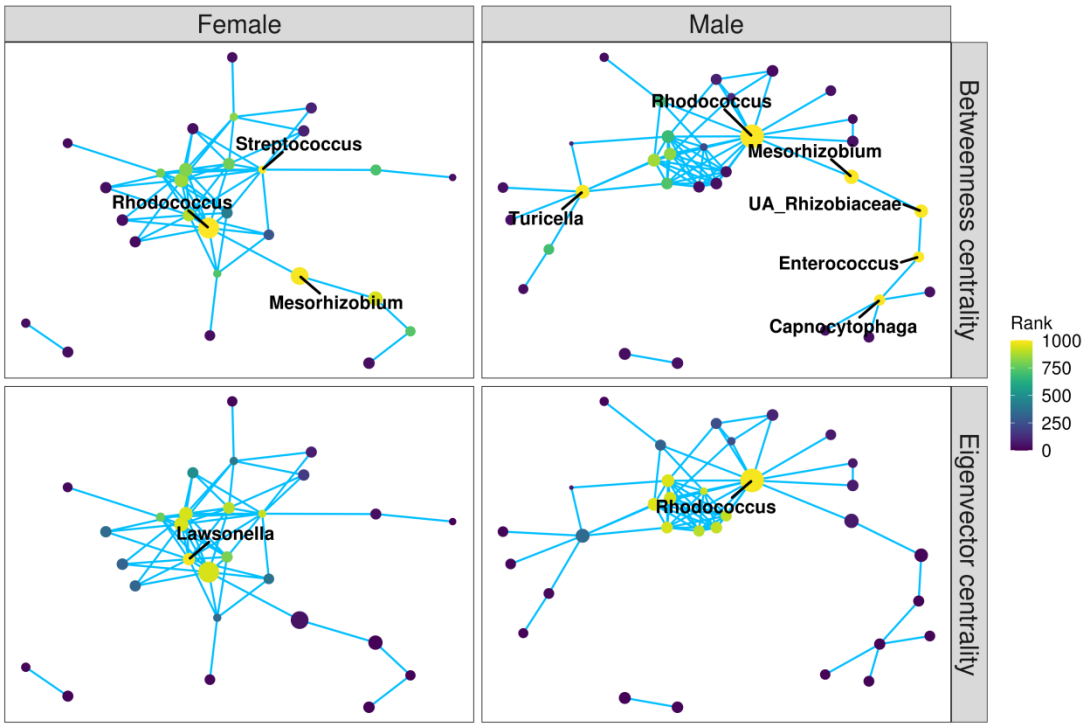

Figure 3.9: Graph representation of each genus for each network in networks of spruce budworm gut tissues from larvae of different sexes, where nodes are shaded depending on their ranks in the null model distribution and their size is proportional to relative abundance. Only genera outside the 95% confidence interval are labeled.

3.6 Discussion

In this study, we used ecological network analysis to quantify and compare microbial co-occurrence in the budworm microbiome based on data sets of bacterial communities from budworm collected from sites across Quebec. We aimed to characterize how these networks differed between whole larvae and gut tissue as well as between host tree species and larval sex, in terms of their structure and taxonomy.

3.6.1 Whole larvae and larval gut tissues

Despite our prediction that co-occurrences (edges) present in the gut tissue should also be present in whole larvae networks due to the spatially nested nature of gut tissues within whole larvae, only one negative co-occurrence was common to both networks (between *Acidiphilium* and *Lawsonella*). Both of these genera inhabit the budworm gut tissue, as *Acidiphilium* are extremophiles thriving in acidic environments and *Lawsonella* was found in gut tissues of fish fed with insect-based diet (L. Li et al., 2020; Panteli et al., 2021). Interestingly, compositional analysis also found that gut tissues and whole larvae had few bacterial taxa in common (Landry et al., Chapter 2).

As predicted, network structure and hub taxa differed between co-occurrence networks for communities of bacteria coming from budworms gut tissues or whole larvae. Network density was much higher in gut tissue networks compared with whole larvae networks indicating that whole larvae were much richer taxonomically for a similar number of co-occurrences. This could be due to the greater selectivity of the lepidopteran gut tissue (Engel & Moran, 2013), fostering more competition in an environment with limited resources and space. This result goes according to our prediction that whole larvae networks would be more diverse taxonomically, because of the selective nature of the insect gut tissue and the exposition of the epidermis to bacterial dispersal from the air (Engel & Moran, 2013). Contrary to our prediction that the diameter of the whole larvae network would be lower than the diameter of the gut tissues network, we observed a higher diameter in the network of whole larvae. It has been hypothesized that a network's diameter was indicative of resistance to perturbation (Delmas et al., 2019), and if this is the case then our results suggest that gut tissue communities are highly sensitive to perturbations, but this remains to be demonstrated empirically.

One of the two genera dominating gut tissue communities, *Rhodococcus*, was also one of the two hub taxa in gut tissue networks in terms of both betweenness centrality and eigenvector centrality. The other dominant genus in gut tissues communities, *Mesorhizobium* was not a hub despite its abundance, showing that measures of network inference are not driven purely by bacterial abundance. *Mesorhizobium*, which represents more than 22% of the bacterial communities in budworm gut tissues (Landry et al., Chapter 2), was not identified as hub taxa, which shows the value of network analysis in assessing the ecological meaning of abundance in a community. In a typical compositional analysis, influence in the community would have been assigned to more abundant taxa such as *Mesorhizobium*, while the influence of rarer taxa, some which we identified as hub taxa (*Bradyrhizobium*, *Streptococcus*) would have been underestimated.

3.6.2 Host tree species variation

Differences in network structure and hub taxa between host tree species were mostly consistent between the whole larvae and gut tissue communities (Figures 3.4 and 3.5). In fact, whole larvae and gut tissue networks varied strongly among host tree species independent of the physiological structure they came from. In both whole larvae and gut tissues networks, microbial co-occurrence networks were more densely connected in microbial communities of larvae found on balsam fir (Figure 3.6). Black spruce networks were generally less central and showed fewer motifs than balsam fir, both in gut tissues and whole larvae. In addition to bacterial community composition of budworm larvae being different across host tree species (Landry et al., Chapter 2), network structure varying along host tree species shows that important processes driving bacterial communities and how they co-occur are linked to the species of tree budworms are feeding on. We speculate that these differences can be attributed to differences in foliage chemistry (Pang, 2010) and to variation in the bacterial

communities associated with foliage of each respective species (Laforest-Lapointe et al., 2016, Landry et al., Chapter 1).

3.6.3 Larval sex variation

In terms of network structure, microbial communities from larvae of different sexes were very similar with *Rhodococcus* being the central hub taxon in both networks, as predicted. However, the identity of network hub taxa other than *Rhodococcus* was different between the sexes, with *Lawsonella*, *Streptococcus*, *Turicella*, *Enterococcus*, *Rhizobiaceae* and *Capnocytophaga* being hub taxa only in the network of one specific sex (Figure 3.9). Landry et al., (Chapter 2), who used the same data for compositional analysis, found a completely different set of sex-specific bacterial genera (*Corynebacterium*, *Staphylococcus*, *Gordonia*, *Actinomyces*, *Leptotrichia*, *Cutibacterium* and *Pediococcus*) than we found. Using network analysis, we were able to characterize bacterial co-occurrences that were not identified using traditional community-focused methods for measuring diversity (Landry et al., Chapter 2). Differences in bacterial co-occurrence networks between larval sexes that were not captured by standard compositional analysis hint at sex-specific microbial associations, which could be related to sex differences in metabolism, diet or behavior in the budworm. Sex-specific differences in behavior and diet have been reported in multiple insect species (Baker et al., 1987; Fanson et al., 2013; Nakashima & Hirose, 2003; Willink & Svensson, 2017), and could be responsible for sex-specific bacterial communities and co-occurrence networks, which have been found in the digestive system of another insect, the fruit fly (Han et al., 2017). Additionally, compared to all other networks, the male gut tissues network showed a notably higher proportion of positive co-occurrence. Exogenous stress impacting whole larvae communities could explain this difference in proportion of positive/negative co-occurrence, as it was found that the more stressed the network, the less it involved negative co-occurrences (Hernandez et al., 2021). Testing

empirically how interactions in bacterial communities of male budworm gut tissues differ from other budworm-associated communities and their relationship to stress would be a promising avenue of research to deepen our knowledge of the link between insect sex and microbiomes.

While it is tempting to infer interactions from the instances of co-occurrence we identified in networks of bacteria associated with the spruce budworm, we believe that more empirical studies are needed to establish unquestionably the link between co-occurrence and ecological interaction. Indeed, despite the fact that many studies assume that microbial co-occurrence networks are evidence of actual microbe interactions, it has been argued that such assumptions are problematic (Blanchet et al., 2020); for example, co-occurrence in samples could be driven by underlying environmental gradients as well as by both direct and indirect interactions in the community, and it is impossible to distinguish these possibilities based on co-occurrence data alone. For microbial communities in particular the spatial scales over which microbes typically interact are many orders of magnitude smaller than the spatial scale of samples, making it very difficult to infer interaction from co-occurrence in these samples. Despite these limitations, our study generated testable hypotheses about the nature of bacterial co-occurrence networks that should be investigated by directly studying interactions among microbial taxa and their hosts in field or lab experiments. Our study is a first step towards integrating bacterial communities into our understanding of the spruce budworm's foodweb and its impact on budworm population dynamics.

3.7 Conclusions

In this study, we characterized co-occurrence patterns in bacterial communities associated with the spruce budworm. Only one co-occurrence was common to both the gut tissues network and the whole larvae network, despite the fact that gut tissues

are physically nested within whole larvae. The co-occurrence network of gut tissues was twice as dense and implicated a less diverse spectrum of bacterial genera than the whole larvae network, as we had predicted since the insect gut is assumed to be a highly selective habitat for microbes (Engel & Moran, 2013). One dominant genus in gut tissue communities, *Rhodococcus*, was identified as hub taxon, although other very abundant bacterial genera were not hub taxa, demonstrating that bacterial abundance does not necessarily equate to influence in co-occurrence networks. Similarly to what was shown in budworm bacterial community relative abundance, host tree species had an important effect on co-occurrence network structure. This effect was similar whether we studied co-occurrence networks of gut tissues and whole larvae, despite our prediction that host tree species would have a stronger effect on the co-occurrence network of whole larvae bacterial communities. Despite Landry et al. (Chapter 2) showing that abundant taxa in budworm gut tissue communities did not vary among larval sexes, we found multiple hub taxa in the co-occurrence networks of these communities that did vary with insect sex. This study generates multiple hypotheses that merit further empirical testing to deepen our understanding of the microbial side of the foodwebs associated with the spruce budworm.

3.8 Acknowledgements

We thank Rob Johns, Paul Whalen, Jérémie Brisson and Marie-Line Buteau for contributing to sampling. We also thank Geneviève Bourret for assistance with sample processing and sequencing.

CONCLUSION

Malgré l'intérêt récent pour l'étude du lien entre les insectes et leur microbiome, on en sait très peu sur les facteurs biologiques et abiotiques qui déterminent la nature des communautés bactériennes associées aux insectes. Cette thèse présente une étude combinant un échantillonnage à très grande échelle spatiale, des méthodes d'analyses écologiques modernes et la science des réseaux pour combler ce manque. À notre connaissance, c'est l'examen le plus exhaustif des facteurs gouvernant l'assemblage et les co-occurrences d'un microbiome d'un Lépidoptère important écologiquement et économiquement, et ce projet contribue aux connaissances disponibles pour comprendre le système dans lequel la TBE évolue. Plus précisément, c'est la première étude mesurant le degré d'impact de communautés bactériennes associées à l'alimentation sur le microbiome d'un Lépidoptère dans un contexte naturel.

Ce projet de doctorat visait à mieux comprendre le lien entre la TBE, son microbiome et son environnement pour élucider comment les microbiomes de Lépidoptères s'assemblent. Plus précisément, l'objectif principal était de caractériser la composition et les co-occurrences de communautés bactériennes associées aux larves de TBE et leur intestin, en plus de mesurer leur variation en fonction de l'emplacement géographique, le climat, leur alimentation, l'espèce de leur arbre hôte et leur sexe. Cette thèse était divisée en trois chapitres, chacun présentant un différent projet contribuant à accomplir cet objectif. Le premier projet consistait en une étude sur une très grande échelle spatiale caractérisant la variation du microbiome de la TBE. Grâce à l'utilisation innovatrice de « Moran Eigenvector Maps » (MEMs) et de la composition des communautés bactériennes du feuillage comme variables

prédictives nous avons pu discriminer l'effet de la variation spatiale, l'alimentation, l'arbre-hôte et le climat sur le microbiome de la TBE. Le deuxième projet a mesuré l'impact du sexe larvaire sur plusieurs aspects du microbiome de l'intestin de la TBE et comment cet impact varie en fonction de l'emplacement géographique. L'importance des communautés de l'intestin vis-à-vis le microbiome de l'insecte en entier a aussi pu être établie en comparant les communautés caractérisées lors du premier projet avec les microbiomes d'intestins caractérisés lors du second. Le troisième projet a permis d'avoir un aperçu des co-occurrences entre microbes associés à la TBE en étudiant la corrélation entre l'abondance de chaque taxon en fonction des autres. En appliquant plusieurs préceptes de la science des réseaux, ceci a permis de mieux comprendre comment ces microbes pourraient affecter leur communautés respectives et comment les relations entre microbes évoluent en fonction des facteurs testés dans les deux premiers chapitres.

4.1 Principaux facteurs expliquant la variation spatiale du microbiome de la TBE

L'objectif du premier chapitre était de déterminer quels étaient les principaux facteurs expliquant la variation du microbiome de la TBE à travers l'aire touchée par l'épidémie. Nous avons démontré que l'assemblage des communautés bactériennes de la TBE est déterminé principalement par la structure spatiale entre les sites, par les communautés bactériennes sur le feuillage et par l'espèce d'arbre hôte. Alors que l'emplacement géographique est un indicateur important de la composition microbiome de la TBE, aucune des variables environnementales que nous avons mesurées ne peut expliquer la variation entre les sites. La majeure partie de la variation parmi les communautés bactériennes de TBE sont reflétées dans les communautés de feuillage qu'elles mangeaient, démontrant que les communautés bactériennes du feuillage sont en fait le principal déterminant de la composition du microbiome de la TBE. La présence de ce lien important entre les communautés bactériennes de larves de la TBE et le feuillage qu'elles consomment soutient

l'hypothèse selon laquelle les microbiomes des lépidoptères sont composés principalement de microbes non résidents et provenant de l'alimentation de l'insecte (Hammer et al., 2017). Cependant, les communautés bactériennes de larves de la TBE et du feuillage sont assez différentes pour suggérer que d'autres processus, en plus du transfert à partir de l'alimentation, jouent un rôle dans l'assemblage du microbiome de la TBE. Outre les communautés foliaires, le meilleur prédicteur de la variation entre les sites est la structure spatiale, représentée par les MEMs, qui représentent une variation spatiale ou environnementale non mesurée. Nous avons également déterminé que l'espèce d'arbre hôte est un déterminant important à la fois de la composition du microbiome de la TBE et de la sensibilité de ce microbiome à l'environnement, parce que les communautés bactériennes de TBE recueillies sur sapin baumier étaient beaucoup plus fortement corrélées avec nos prédicteurs spatiaux et environnementaux. Grâce à ces résultats, nous comprenons maintenant mieux les facteurs qui déterminent en milieu naturel l'assemblage et la biogéographie du microbiome de la TBE.

4.2 Variation entre les communautés bactériennes de la TBE en fonction du sexe et entre les communautés d'intestins et de larves entières

L'objectif du deuxième chapitre était de caractériser les différences entre les communautés bactériennes d'intestins de larves mâles, d'intestins de larves femelles et de larves entières. Nous avions prédit la présence de différences de composition et de diversité entre les microbiomes d'intestins de larves mâles et femelles. Les communautés bactériennes d'intestins de larves femelles étaient plus diverses que celles de larves mâles. En termes de composition, il n'y avait pas de différences majeures entre ces communautés, à l'exception de quelques taxons différentiellement abondants qui étaient associés à un sexe en particulier. En revanche, nous avons détecté d'importantes différences de composition entre les communautés bactériennes des intestins de TBE par rapport aux communautés de larves entières. En général, les

communautés intestinales étaient beaucoup plus sensibles à l'emplacement géographique de la larve et étaient fortement dominées par deux taxons, *Rhodococcus* et *Mesorhizobium,* dont l'abondance dépendait fortement du site. Ce résultat indique qu'une partie non négligeable des communautés bactériennes de la TBE vit à l'extérieur de leurs intestins, et que le choix de la partie du corps à échantillonner a des implications importantes pour les résultats des études sur le microbiome des Lépidoptères. Ces résultats soutiennent aussi l'hypothèse que l'intestin de la TBE est un milieu sélectif limitant l'établissement d'une bonne proportion des bactéries présentes dans l'environnement et l'alimentation de la TBE.

4.3 Les co-occurrences des communautés bactériennes de la TBE et leur variation

Nous avons mesuré les patrons de co-occurrence dans les communautés bactériennes associées à la TBE et comment elles varient entre les intestins et les larves entières, les espèces d'arbres hôtes et les sexes larvaires. Nous avons découvert que les réseaux de co-occurrence bactérienne des tissus intestinaux des larves étaient organisés autour de *Rhodococcus* en tant que taxon central, et que les communautés de tissus intestinaux étaient impliquées dans des réseaux beaucoup plus denses que les réseaux de larves entières. Les communautés bactériennes de l'intestin de la TBE sont donc beaucoup moins diverses taxonomiquement mais à la fois aussi riches en co-occurrences microbiennes que le reste de l'insecte. Les réseaux de taxons bactériens provenant d'insectes récoltés sur le sapin baumier étaient en général plus centraux, denses et complexes que les réseaux de bactéries provenant d'insectes récoltés sur l'épinette noire, qu'ils proviennent de communauté d'intestins ou de larves entières. Cette variation importante de réseaux de co-occurrence indique que l'arbre hôte a un impact important sur les communautés bactériennes de la TBE, que ce soit à cause d'une différence de composition chimique ou de composition de communautés bactériennes associées au feuillage de chaque espèce. Les réseaux de différents sexes larvaires étaient similaires en termes de structure de réseau et de

centralité, et nous avons pu identifier plusieurs taxons moyeux spécifiques au sexe qui n'ont pas été identifiés comme tel lors de l'analyse par abondance relative lors du chapitre 2. Ceci indique que l'analyse de réseau de co-occurrence est complémentaire à l'analyse par abondance relative pour identifier quels taxons peuvent avoir un impact important sur leur communauté et leur hôte. Caractériser ces réseaux a permis de comprendre quels facteurs entraînent la variation de la structure et des taxons moyeux des réseaux de cooccurrence bactérienne chez la TBE. À notre avis, la quantification du réseau de cooccurrence bactérienne de la TBE était la première étape nécessaire pour comprendre le réseau trophique microbien de la TBE, ce qui contribue à combler une lacune de connaissances clé par rapport à notre compréhension du réseau trophique animal et végétal de l'insecte. Malgré que nous ne pouvons assumer que les patrons de co-occurrence observés représentent bel et bien des interactions entre microbes (Blanchet et al., 2020), ces patrons que nous avons mesurés et leur variation contribuent à générer des hypothèses relatives à la façon dont les bactéries interagissent entre elles et avec leur hôte. Ces hypothèses pourraient être testées en mesurant directement les interactions entre bactéries chez la TBE pour voir si les patrons de co-occurrence observés prédisent effectivement des interactions écologiques.

4.4 Limites du projet

Malgré le fait que ce projet a permit d'accomplir les objectifs que nous avions fixés, certains aspects laissent place à l'amélioration:

1) L'analyse de partitionnement de la variance utilisée lors du chapitre 1 n'a pu expliquer qu'une petite proportion de la variance présente à travers nos échantillons (16%) avec les facteurs inclus dans notre modèle. Malgré qu'un total de variance expliquée très bas est commun lorsque ce type d'analyse est utilisé dans un contexte de communautés bactériennes (Li et al., 2020; Logares et al., 2020; Yan et al., 2016),

ceci indique qu'un grande partie de la variation présente dans les communautés de microbes associées à la TBE n'était pas expliquée par les variables prédictives utilisées. Le faible pouvoir explicatif de nos variables mesurées pourrait être dû à l'importance des processus stochastiques dans l'assemblage et la biogéographie de la communauté du microbiome de la TBE, tels que la dispersion des insectes hôtes et des microbes. La variation génétique de la TBE ou des arbres hôtes pourrait être une explication potentielle cette variation inexpliquée. Par contre, la structure génétique spatiale de la TBE dans notre zone d'étude était généralement homogène, avec quelques différences génétiques subtiles entre les deux régions (Larroque et al., 2020). Étant donné que la majeure partie de la variation inter-site des communautés bactériennes était à une échelle spatiale plus fine que celle de ces différences régionales de structure génétique chez la TBE, il semble probable que l'impact de la variation génétique chez la TBE sur les communautés bactériennes soit minime. Cependant, comme les communautés foliaires sont étroitement liées aux communautés larvaires, lier la variation génétique de l'arbre hôte dans la zone d'épidémie aux communautés bactériennes de la TBE pourrait être une voie prometteuse pour la recherche future. Par exemple, les conifères produisent des tannins antibactériens dans leur feuillage (Pang, 2010), et des différences génétiques dans la production de ces molécules pourraient effectivement mener à un variation dans les communautés associées au feuillage et par le fait même dans les communautés associées à la TBE.

2) Dans le contexte de comparaison entre les communautés et les réseaux de larves entières et d'intestin, les deux ensembles de données provenaient d'échantillons collectés pendant deux années différentes qui ont été amplifiés et séquencés séparément (une source de biais reconnue (Pollock et al., 2018)). Conséquemment, les différences entre ces deux groupes ont possiblement été influencées par des différences interannuelles ou entre instances de séquençage. Cependant, les échantillons des deux groupes ont été traités selon les mêmes protocoles, par les

mêmes personnes et leurs jeux de données respectifs ont été intégrés lors des toutes premières étapes du traitement bioinformatique pour minimiser la variation due à la différence de type d'échantillon. Aussi, les études à grande échelle temporelle du microbiome de systèmes naturels démontrent très peu de variation interannuelle (Cáliz et al., 2018; Epstein et al., 2019) et le temps est un facteur très mineur dans l'assemblage de microbiome du feuillage d'arbres feuillus au Québec (Laforest-Lapointe et al., 2016). Il est donc fort probable que les importantes différences observées entre les communautés d'intestins et de larves entières soient en majeure partie dues à la variation de l'habitat dans l'insecte qui est échantillonné, comme il a été démontré chez d'autres insectes et mammifères (Dekaboruah et al., 2020; Otti, 2015; Stehlikova et al., 2019).

3) Comme la nature de nos analyses se limitait à la caractérisation d'abondance relative de taxons bactériens, il est impossible de déterminer si les différents facteurs étudiés avaient un impact sur la quantité de bactéries présentes dans la TBE et son intestin. Il serait donc pertinent de faire suite à ce projet en réévaluant l'impact du sexe, de la situation géographique et de l'arbre hôte sur à la fois la composition mais aussi l'abondance des communautés associées à la TBE. On pourrait ainsi déterminer si des processus menant à une diminution de la quantité de bactéries sont à l'œuvre dans une espèce d'arbre hôte ou un sexe larvaire. De plus, on pourrait confirmer si l'effet sélectif des intestins de la tordeuse limite bel et bien l'établissement d'une population abondante bactérienne en plus de minimiser l'étendue des taxons bactériens pouvant s'y établir.

4) Nous pensons que davantage d'études empiriques sont nécessaires pour inférer des interactions écologiques à partir des patrons de co-occurrence que nous avons observé. En effet, malgré le fait que de nombreuses études supposent que les réseaux de co-occurrence microbienne sont la preuve d'interactions microbiennes réelles, il a été avancé que la complexité des interactions écologiques entre microbes par rapport à

nos capacité d'échantillonnage rendent de telles assomptions problématiques (Blanchet et al., 2020). Plus précisément, la co-occurrence dans les échantillons pourrait être due à des gradients environnementaux sous-jacents ainsi qu'à des interactions directes et indirectes dans la communauté, et il est impossible de distinguer ces possibilités sur la base de seulement données de co-occurrence. Pour les communautés microbiennes en particulier, l'échelle spatiale sur laquelle les microbes interagissent est généralement de plusieurs ordres de grandeur plus petite que l'échelle spatiale des échantillons, ce qui rend très difficile de déduire l'interaction à partir de la co-occurrence dans ces échantillons. Notre étude génère néanmoins des hypothèses intéressantes sur la nature des réseaux de co-occurrence bactérienne qui devraient être testées en étudiant directement les interactions entre les taxons microbiens et leurs hôtes dans des expériences de terrain ou de laboratoire. Dans cette perspective, notre étude est la première étape pour comprendre comment les bactéries associées à la TBE interagissent entre elles et avec leur hôte. Observer la présence de ces interactions empiriquement permettrait d'intégrer les communautés bactériennes dans notre compréhension du réseau trophique de la TBE et de son impact sur la dynamique despopulations de l'insecte.

4.5 Recommandations pour le futur

Malgré les limites de cette thèse, ce projet contribue grandement à notre compréhension des communautés bactériennes associées à la TBE et leurs patrons de co-occurrence en plus de caractériser leur variation en fonction de plusieurs facteurs parmi lesquels l'emplacement géographique, l'espèce d'arbre-hôte et le sexe de l'insecte. Nous proposons trois directions que les prochaines études pourraient prendre pour capitaliser sur les connaissances acquises suite à ce projet.

1) Les résultats présentés dans cette thèse révèlent un lien important entre les communautés bactériennes sur le feuillage consommé par la TBE et les communautés

de larves entières de la TBE. Par contre, certaines différences existaient quand même entre les deux types de communautés. En lien avec la théorie selon laquelle les microbiomes d'intestins de Lépidoptères proviennent principalement de l'alimentation de l'insecte (Hammer et al., 2017), nous pensons qu'il est important de mesurer directement les différences de composition et d'abondance totale entre le microbiome de l'intestin de la TBE et celui du feuillage, pour mieux comprendre comment le transfert de bactéries par la diète est accomplit chez la tordeuse. On pourrait ainsi détecter si l'intestin est bel et bien un filtre environnemental où seule une proportion de taxons ou d'individus bactériens peut s'y établir. On pourrait aussi mesurer le pH ou l'activité de gènes modifiant le pH au cours de la vie de l'insecte tout en observant le transfert par la diète afin de déterminer si l'insecte influence lui-même directement son microbiome intestinal. À cette expérience pourrait s'ajouter aussi la détection de tannins antibactériens présents dans le feuillage de conifère (Pang, 2010), ou mesurer l'activité de gènes de défense antibactériens dans le feuillage et mesurer leur impact sur l'abondance et la composition des communautés transférées à l'insecte par la diète. La variation de tannins et autres molécules antibactériennes présents dans le feuillage pourraient expliquer l'importante variabilité spatiale de la composition du microbiome observée lors du chapitre 1 et qui demeure inexpliquée par les facteurs spatiaux et environnementaux que nous avons mesurés.

2) Le troisième chapitre de cette thèse a permit d'identifier un grand nombre de co-occurrences et de taxons moyeux parmi les genres bactériens formant le microbiome de la TBE. Par contre, comme les méthodes de mesure de co-occurrences ne détectent pas forcément les interactions écologiques derrière ces patrons de co-occurrences (Blanchet et al., 2020), nous ne pouvons déterminer si ces genres bactériens interagissent réellement dans le contexte du microbiome de la TBE et si les taxons moyeux identifiés ont réellement une grande influence sur leur communauté et/ou leur hôte (Agler et al., 2016). Pour combler à ce manque, les taxons bactériens

impliqués dans les co-occurrences les plus intéressantes ou ceux identifiés comme taxons moyeux pourraient être cultivés en laboratoire. On pourrait tester leur interaction avec d'autres taxons bactériens en les cultivant ensemble et en observant l'effet sur la croissance. On pourrait aussi directement infecter le corps ou l'alimentation de la TBE avec une de ces bactéries ou une combinaison de celles-ci, et mesurer leur effet sur le poids et la santé de la chenille. Ces tests permettraient à la fois de confirmer à quel point les méthodes de mesure de co-occurrence identifient bel et bien des interactions écologiques, et d'identifier des taxons bactériens ayant un effet bénéfique ou délétère sur la santé de la TBE.

3) Cette thèse apporte une grande contribution pour prédire quels types de communautés ou de co-occurrences bactériennes sont présents dans le microbiome de la TBE en fonction de plusieurs facteurs spatiaux, environnementaux, et biologiques. Par contre, nous croyons que le chaînon manquant pour pouvoir relier le microbiome aux dynamiques de population et au contrôle de la TBE demeure l'effet des communautés bactériennes ou de taxons particuliers sur l'insecte lui-même. Dans le cas où nous pourrions identifier des taxons ou des groupes de bactéries qui interagissent et ont un effet sur la santé de la tordeuse, il serait ensuite possible de mesurer à long terme la présence de ces groupes bactériens chez les populations de la TBE. Ceci permettrait potentiellement d'identifier quels membres du microbiome pourraient avoir un impact sur le déclenchement d'épidémies de TBE, de façon similaire à nos connaissances sur le réseau trophique animal dont la TBE fait partie et son impact sur les dynamiques de population de la TBE (Eveleigh et al., 2007; Marrec et al., 2018). On pourrait ensuite relier les connaissances acquises lors de cette thèse pour identifier quel emplacement géographique et quelle espèce d'arbre hôte pourrait potentiellement être l'épicentre d'une nouvelle épidémie. À plus court terme, comme le microbiome de la TBE pourrait avoir un impact sur l'efficacité de la formulation *Bacillus thurigiensis* utilisée pour contrôler les population de TBE (Broderick et al., 2009; Frankenhuyzen et al., 2010), connaître les agents pathogènes

potentiels de la TBE pourrait contribuer à l'utilisation optimale de cet insecticide. On pourrait identifier les peuplements donc qui bénéficieraient plus de l'application de l'insecticide car ils sont riches en pathogènes qui tueraient l'insecte suite à la perforation de son intestin par *B. thurigiensis*. On pourrait aussi supplémenter la formulation de l'insecticide avec un pathogène spécifique à la TBE, garantissant ainsi son efficacité peu importe le microbiome de la TBE touchée.

Dans l'ensemble, notre étude a quantifié les facteurs à l'origine de la variation du microbiome de la TBE et identifié des pistes prometteuses pour de futures études et pour le contrôle de cet insecte ravageur. La disponibilité croissante d'approches de séquençage pour quantifier la diversité microbienne a ouvert les portes à une meilleure compréhension du lien entre la TBE et son microbiome, et nous permettra éventuellement de quantifier l'ensemble du réseau trophique de cet insecte ravageur forestier important et mieux comprendre la dynamique de ses populations.

ANNEXE A

FIGURE SUPPLÉMENTAIRE 1.1

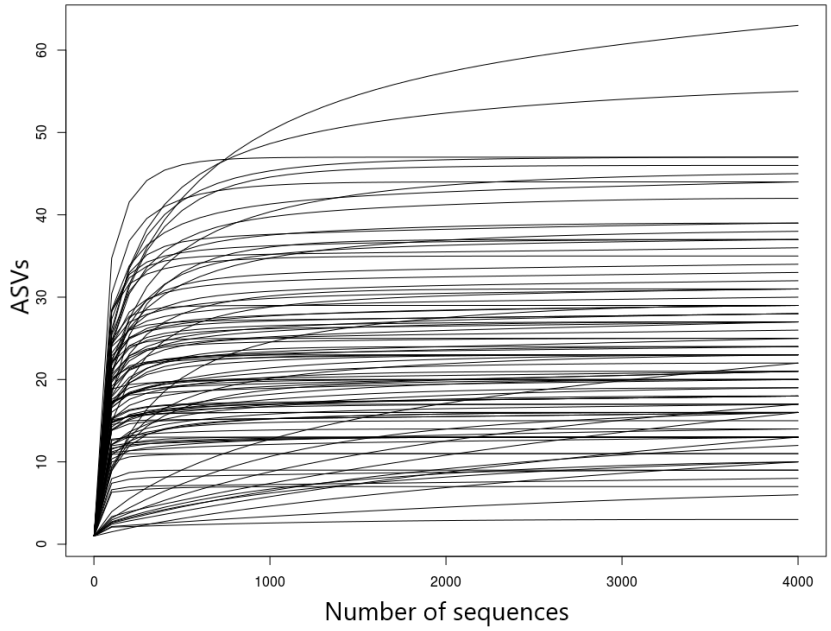

Figure Supplémentaire 1.1: Courbe de raréfaction du nombre d'ASV en fonction du nombre de séquences par échantillon. Les courbes de la plupart des échantillons ont atteint un plateau à une profondeur de séquençage de moins de 1000 séquences par échantillon, démontrant que le seuil de 4000 séquences par échantillon utilisé dans cette étude était suffisant pour capturer de manière adéquate la diversité bactérienne dans ces échantillons.

ANNEXE B

FIGURE SUPPLÉMENTAIRE 1.2

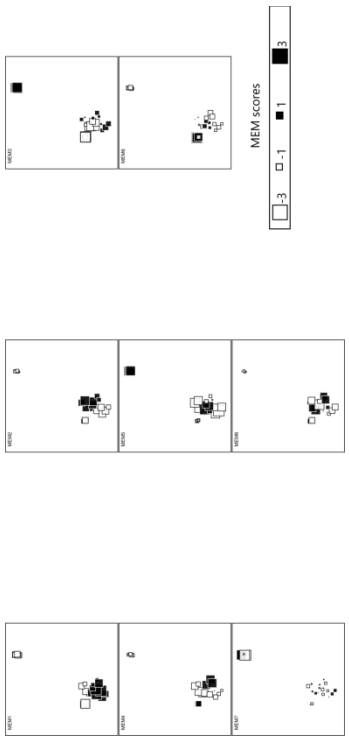

Figure Supplémentaire 1.2: Les scores MEM superposés à la position géographique approximative des sites d'échantillonnage.

ANNEXE C

FIGURE SUPPLÉMENTAIRE 1.3

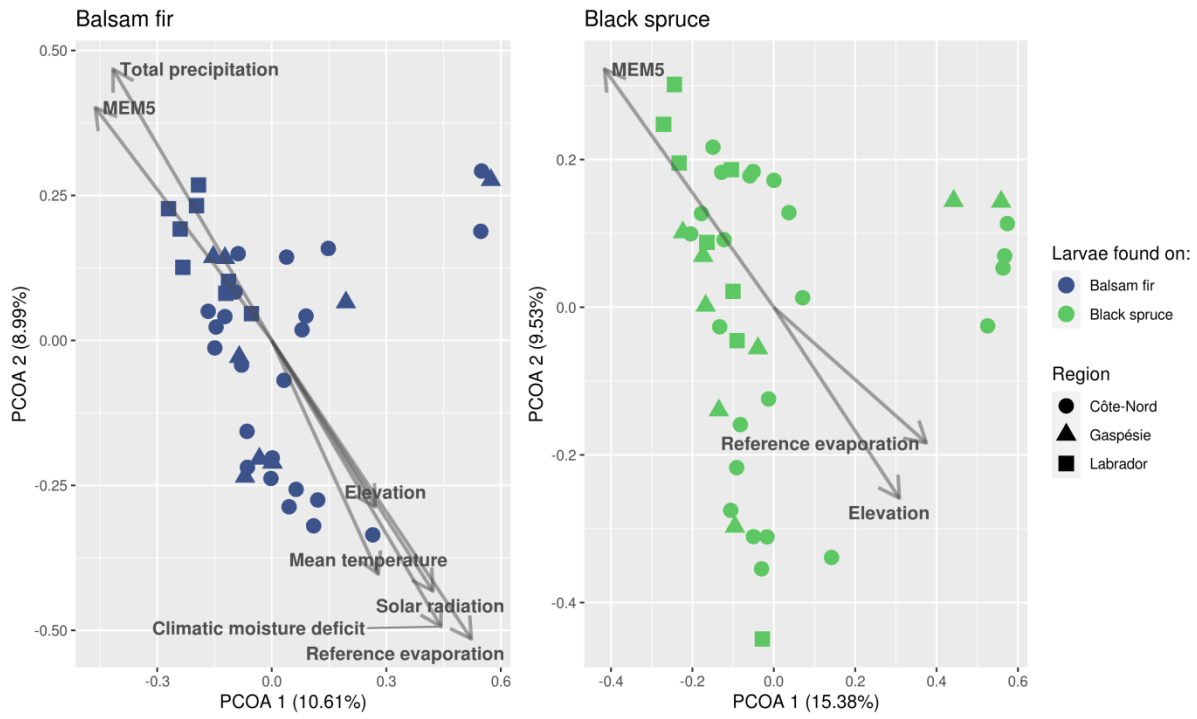

Figure Supplémentaire 1.3: Ordination PCoA sur la dissimilarité Bray-Curtis avec une corrélation significative des vecteurs en ajustant les facteurs mesurés sur l'espace d'ordination pour les communautés bactériennes de la TBE prélevée sur le sapin baumier par rapport aux épinettes noires. Les facteurs indiqués avaient une corrélation significative avec les axes d'ordination ($P < 0,05$).

ANNEXE D

TABLEAU SUPPLÉMENTAIRE 1.1

Tableau supplémentaire 1.1: Variables spatiales et environnementales à 26 sites au Québec où la composition de la communauté bactérienne de la TBE et du feuillage a été échantillonnée.

ID	Date	Latitude	Longitude	Elevation (m)	Mean temperature (^{o}C)	Total precipitations (mm)	Solar radiation (MJ/m^{2}*d)	Reference evaporation (mm)	Climatic moisture deficit (mm)	Relative humidity (%)
QC1	18/06/2015	48.19	-67.26	281	15.2	79	20.6	123	44	60
QC2	18/06/2015	48.29	-67.42	265	15.1	78	20.5	122	44	60
QC3	18/06/2015	48.27	-67.44	226	15.2	77	20.6	122	45	60
QT2	18/06/2015	48.28	-67.47	363	14.8	79	20.7	120	41	60
QT3	18/06/2015	48.24	-67.46	295	15.1	78	20.8	122	44	60
QT1	24/06/2015	48.19	-67.27	276	15.3	79	20.4	124	45	60
NL1	30/06/2015	53.31	-60.17	0	12.9	139	17.7	108	0	60
NL2	03/07/2015	53.39	-60.41	10	13.2	137	17.7	111	0	59
NL3	03/07/2015	53.34	-60.20	0	12.9	140	17.7	109	0	60
NL4	03/07/2015	53.28	-60.49	1	13.3	131	17.7	111	0	59
T001	05/07/2015	48.96	-69.19	183	15.3	76	19.9	121	45	61
T002	05/07/2015	49.28	-69.53	388	15.3	85	20.1	127	42	57
T003	05/07/2015	49.65	-69.84	548	14.3	92	20	122	30	57
T008	07/07/2015	49.24	-68.60	261	15.6	77	19.7	118	41.7	63
T009	07/07/2015	49.61	-69.08	433	15.7	86	19.4	124	39	59
T007	08/07/2015	49.39	-68.07	218	15.3	74	19.8	116	41.7	63
T010	08/07/2015	49.75	-68.54	300	15.3	83	19.6	121	38.3	60
NL5	09/07/2015	53.28	-60.62	68	14.6	130	17.7	110	0	62
E1	11/07/2015	49.35	-67.83	98	17.5	69	19.7	123	54.3	64
T011X	15/07/2015	49.90	-68.20	145	17.4	77	18.4	125	48.3	62
T012X	15/07/2015	50.23	-68.24	518	16.1	90	18.8	121	31.3	61
T006X	17/07/2015	49.54	-67.49	178	17.4	74	19.1	119	44.7	66
T015	17/07/2015	50.09	-67.05	159	17.4	80	18.8	118	37.7	67
T016	17/07/2015	50.31	-67.40	340	16.7	85	18.6	118	34	64
T013	18/07/2015	50.77	-68.66	416	16.4	91	18.2	125	33.7	59
T017	18/07/2015	50.39	-68.85	550	16.1	91	18.6	124	32.7	60

ANNEXE E

TABLEAU SUPPLÉMENTAIRE 1.2

Tableau supplémentaire 1.2: Corrélation entre les vecteurs ajustés et les axes d'ordination dans l'espace d'ordination de la Figure 1.3. Les facteurs significativement corrélés (P < 0,05) sont indiqués en gras.

Factor	Correlation with PCoA 1	Correlation with PCoA 2	R^2	P-Value
Elevation	**-0.46**	**-0.89**	**0.144**	**0.001**
Mean temperature	**-0.28**	**-0.96**	**0.216**	**0.001**
Total precipitation	**0.45**	**0.89**	**0.188**	**0.001**
Reference evaporation	**-0.49**	**-0.87**	**0.297**	**0.001**
Climatic moisture deficit	**-0.47**	**-0.88**	**0.207**	**0.001**
MEM5	**0.63**	**0.78**	**0.277**	**0.001**
Solar radiation	**-0.67**	**-0.75**	**0.105**	**0.004**
MEM8	**0.88**	**-0.47**	**0.072**	**0.021**
MEM1	**-0.51**	**-0.86**	**0.064**	**0.035**
MEM2	0.90	-0.43	0.043	0.113
MEM3	0.29	0.96	0.036	0.143
Relative humidity	0.86	-0.51	0.020	0.361
Host tree damage	-0.06	1.00	0.018	0.414
MEM4	-0.52	-0.85	0.009	0.624
MEM6	-0.46	-0.89	0.005	0.794
MEM7	-0.41	0.91	0.002	0.877

ANNEXE F

TABLEAU SUPPLÉMENTAIRE 1.3

Tableau supplémentaire 1.3: Corrélation entre les vecteurs ajustés et les axes d'ordination dans l'espace d'ordination de la Figure supplémentaire 1.3. Les facteurs significativement corrélés (P < 0,05) sont indiqués en gras.

	Balsam fir				Black spruce			
Factor	**Correlation with PCoA 1**	**Correlation with PCoA 2**	**R^2**	**P-Value**	**Correlation with PCoA 1**	**Correlation with PCoA 2**	**R^2**	**P-Value**
MEM5	**-0.76**	**0.66**	**0.375**	**0.001**	**-0.79**	**0.61**	**0.277**	**0.004**
Reference evaporation	**0.71**	**-0.70**	**0.538**	**0.001**	**0.90**	**-0.44**	**0.174**	**0.033**
Solar radiation	**0.70**	**-0.72**	**0.365**	**0.001**	1.00	-0.01	0.113	0.127
Climatic moisture deficit	**0.67**	**-0.74**	**0.440**	**0.001**	0.87	-0.50	0.094	0.171
Total precipitation	**-0.66**	**0.75**	**0.392**	**0.001**	-0.80	0.60	0.093	0.183
Mean temperature	**0.57**	**-0.82**	**0.240**	**0.007**	0.30	-0.95	0.129	0.075
Elevation	**0.69**	**-0.73**	**0.156**	**0.049**	**0.77**	**-0.64**	**0.162**	**0.029**
MEM3	-0.57	0.82	0.111	0.126	-0.37	0.93	0.001	0.982
MEM2	-0.81	0.59	0.091	0.156	-0.61	-0.79	0.120	0.086
MEM8	-0.57	-0.82	0.078	0.229	-0.95	-0.30	0.147	0.055
MEM1	0.74	-0.67	0.061	0.369	0.78	-0.62	0.073	0.245
MEM4	0.42	0.91	0.054	0.37	0.13	-0.99	0.011	0.818
Relative humidity	-0.99	0.10	0.044	0.459	-0.61	-0.79	0.120	0.096
Host tree damage	-0.87	0.49	0.020	0.719	0.17	0.99	0.106	0.143
MEM6	-0.74	-0.68	0.001	0.973	0.82	-0.57	0.027	0.599
MEM7	-1.00	-0.02	0.000	0.991	0.24	0.97	0.023	0.654

ANNEXE G

TABLEAU SUPPLÉMENTAIRE 1.4

Tableau supplémentaire 1.4: Corrélation Pearson entre facteurs environnementaux.

	Host tree Damage	Mean temperature	Total precipitation	Solar radiation	Reference evaporation	Climatic moisture deficit	Relative humidity	Degrees of freedom
Host tree Damage	1.00	0.29	-0.41	0.13	-0.11	0.38	0.36	16
Mean temperature	0.29	1.00	0.01	-0.73	-0.01	0.02	0.65	16
Total precipitation	-0.41	0.01	1.00	-0.58	0.24	-0.91	-0.42	16
Solar radiation	0.13	-0.73	-0.58	1.00	0.06	0.59	-0.35	16
Reference evaporation	-0.11	-0.01	0.24	0.06	1.00	0.19	-0.70	16
Climatic moisture deficit	0.38	0.02	-0.91	0.59	0.19	1.00	0.15	16
Relative humidity	0.36	0.65	-0.42	-0.35	-0.70	0.15	1.00	16

ANNEXE H

TABLEAU SUPPLÉMENTAIRE 1.5

Tableau supplémentaire 1.5: Corrélation de Pearson entre MEMs.

	MEM1	MEM2	MEM3	MEM4	MEM5	MEM6	MEM7	MEM8	MEM9	Degrees of freedom
MEM1	1.00	-0.13	0.55	-0.48	0.00	-0.09	0.15	0.20	0.10	16
MEM2	-0.13	1.00	-0.24	0.13	0.03	0.04	-0.25	-0.25	-0.09	16
MEM3	0.55	-0.24	1.00	-0.07	0.62	-0.47	0.49	0.10	0.26	16
MEM4	-0.48	0.13	-0.07	1.00	-0.12	-0.11	-0.53	-0.16	-0.27	16
MEM5	0.00	0.03	0.62	-0.12	1.00	0.00	0.67	-0.33	0.27	16
MEM6	-0.09	0.04	-0.47	-0.11	0.00	1.00	0.05	-0.09	0.16	16
MEM7	0.15	-0.25	0.49	-0.53	0.67	0.05	1.00	0.30	0.80	16
MEM8	0.20	-0.25	0.10	-0.16	-0.33	-0.09	0.30	1.00	0.62	16
MEM9	0.10	-0.09	0.26	-0.27	0.27	0.16	0.80	0.62	1.00	16

ANNEXE I

TABLEAU SUPPLÉMENTAIRE 1.6

Tableau supplémentaire 1.6: Variance expliquée par les prédicteurs dans l'analyse de partitionnement de la variance (Figure 1.4).

Predictor	Adjusted R^2
Environment	0.010
Spatial structure	0.054
Foliage communities	0.093
Host tree species	0
Environment + Foliage communities	0.000
Foliage communities + Host tree species	0.027
Environment + Spatial structure + Host tree species	0.000
Spatial structure + Foliage communities + Host tree species	0.005

ANNEXE J

FIGURE SUPPLÉMENTAIRE 2.1

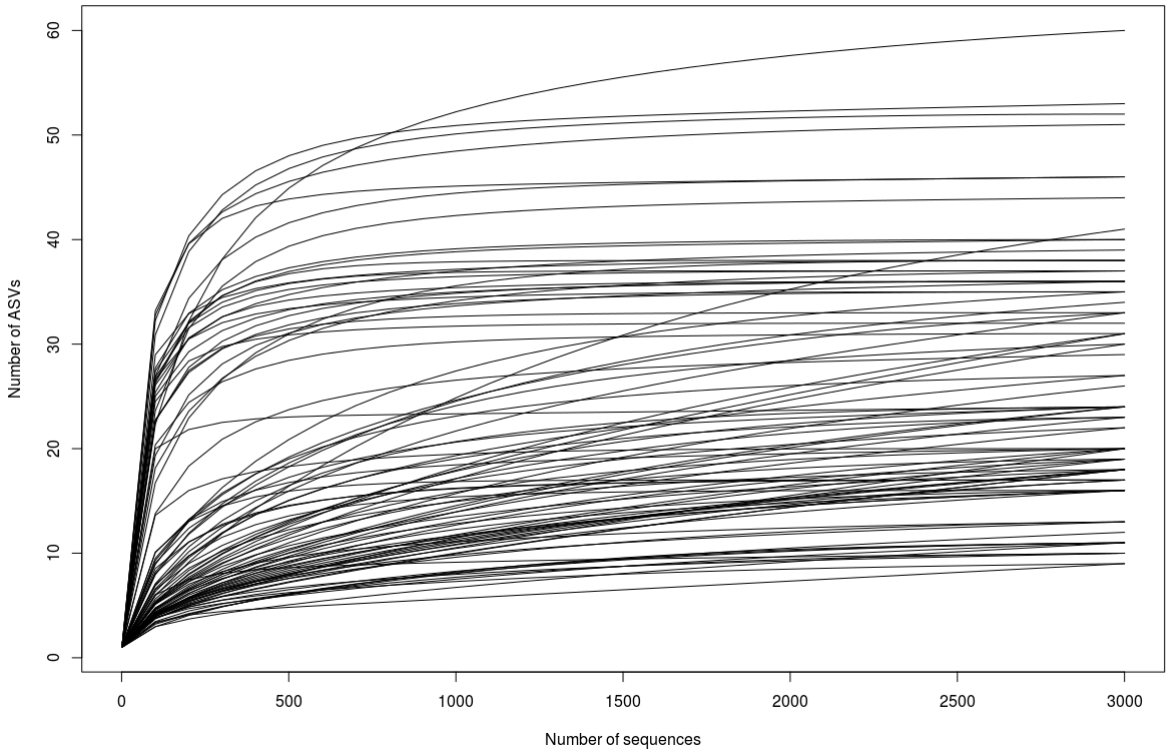

Figure supplémentaire 2.1: Courbe de raréfaction du nombre d'ASVs par rapport au nombre de séquences par échantillon dans le jeu de donnée pour la comparaison entre les communautés des deux sexes larvaires.

ANNEXE K

FIGURE SUPPLÉMENTAIRE 2.2

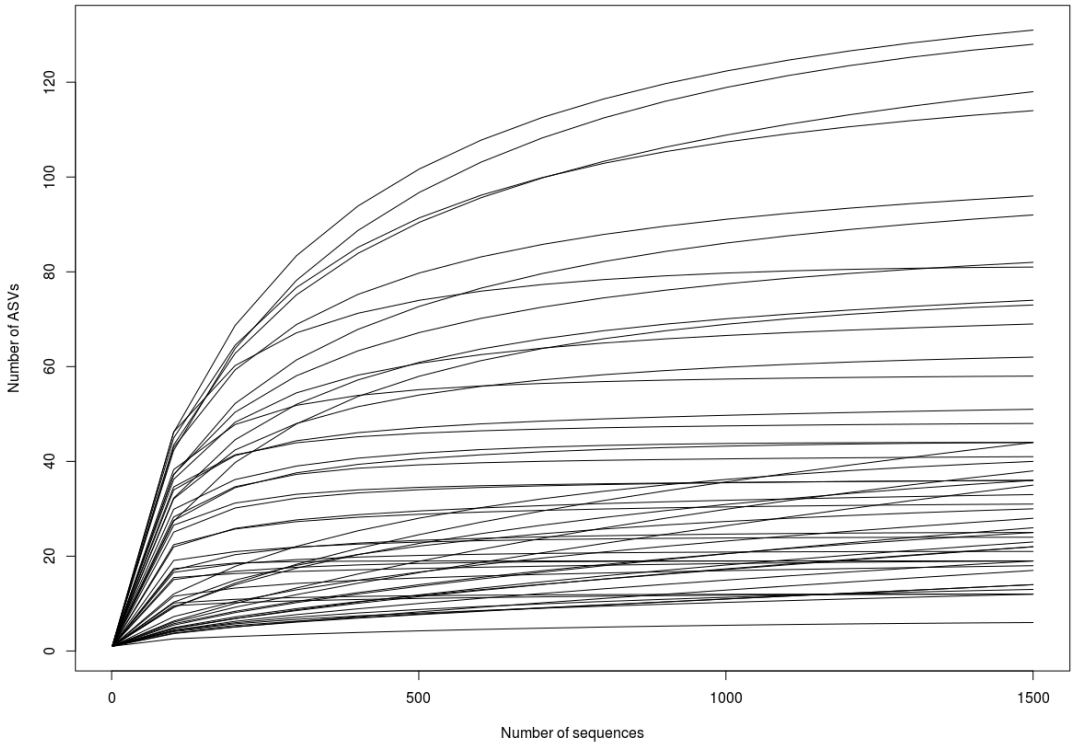

Figure supplémentaire 2.2: Courbe de raréfaction du nombre d'ASVs par rapport au nombre de séquences par échantillon dans le jeu de donnée pour la comparaison entre les communautés de tissus intestinaux et de larves entières.

ANNEXE L

TABLEAU SUPPLÉMENTAIRE 2.1

Tableau supplémentaire 2.1: Variation de la composition des communautés bactériennes de la TBE dans le tissu intestinal en fonction du site d'échantillonnage et du sexe des larves nichées dans le site, testée à l'aide d'un PERMANOVA basé sur la dissimilarité de Bray-Curtis, modélisé séparément pour chaque site.

	Df	Sum Sq	Mean Sq	F-Value	R^2	P-Value
All sites						
Sex	1	0.50	0.50	1.64	0.019	0.168
Residuals	85	25.81	0.30		0.981	
Total	86	26.31			1	
T001						
Sex	1	0.00	0.00	0.45	0.022	0.748
Residuals	20	0.08	0.00		0.978	
Total	21	0.08			1	
T003						
Sex	1	0.00	0.00	1.74	0.073	0.132
Residuals	22	0.02	0.00		0.927	
Total	23	0.03			1	
T006						
Sex	1	0.31	0.31	1.25	0.152	0.131
Residuals	7	1.75	0.25		0.848	
Total	8	2.06			1	
T010						
Sex	1	0.19	0.19	0.65	0.056	0.841
Residuals	11	3.24	0.29		0.944	
Total	12	3.42			1	
T013						
Sex	1	0.02	0.02	1.71	0.091	0.129
Residuals	17	0.25	0.01		0.909	
Total	18	0.27			1	

ANNEXE M

TABLEAU SUPPLÉMENTAIRE 2.2

Tableau supplémentaire 2.2: Analyse discriminante DeSEQ2 des ASVs bactériens différentiellement abondants dans les communautés bactériennes des tissus intestinaux des larves de la TBE de sexes différents, illustrée à la figure 2.5. Seuls les ASVs représentant plus de 0,01 % de l'ensemble de l'analyse sont présentés dans le tableau et dans la figure 2.5.

	Phylum	Class	Order	Family	Genus	log2FoldChange	% relative abundance	Associated with
ASV242	Actinobacteriota	Actinobacteria	Corynebacteriales	Nocardiaceae	Gordonia	-2.07	0.11	Female
ASV369	Fusobacteriota	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichia	-1.92	0.07	Female
ASV93	Firmicutes	Bacilli	Staphylococcales	Staphylococcaceae	Staphylococcus	-1.52	0.06	Female
ASV136	Actinobacteriota	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces	-4.53	0.05	Female
ASV134	Actinobacteriota	Actinobacteria	Propionibacteriales	Propionibacteriaceae	Cutibacterium	-1.94	0.05	Female
ASV243	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-1.61	0.05	Female
ASV260	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-1.82	0.04	Female
ASV503	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Pediococcus	1.29	0.04	Male
ASV397	Actinobacteriota	Actinobacteria	Propionibacteriales	Propionibacteriaceae	Cutibacterium	-2.24	0.04	Female
ASV48	Firmicutes	Bacilli	Staphylococcales	Staphylococcaceae	Staphylococcus	-1.39	0.04	Female
ASV609	Patescibacteria	Saccharimonadia	Saccharimonadales	Saccharimonadaceae	unassigned	2.34	0.03	Male
ASV624	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Capnocytophaga	1.196	0.03	Male
ASV361	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Enhydrobacter	-2.08	0.03	Female
ASV32	Firmicutes	Bacilli	Staphylococcales	Staphylococcaceae	Staphylococcus	-1.58	0.02	Female
ASV824	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium	-1.21	0.02	Female
ASV205	Actinobacteriota	Actinobacteria	Micrococcales	Dermabacteraceae	Dermabacter	1.69	0.02	Male
ASV324	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-1.71	0.02	Female
ASV373	Proteobacteria	Gammaproteobacteria	Burkholderiales	Neisseriaceae	Neisseria	1.12	0.02	Male
ASV800	Actinobacteriota	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces	-1.09	0.02	Female
ASV1226	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium	0.95	0.01	Male
ASV200	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Stenotrophomonas	-1.06	0.01	Female
ASV129	Firmicutes	Bacilli	Staphylococcales	Staphylococcaceae	Staphylococcus	-1.05	0.01	Female
ASV470	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	1.08	0.01	Male
ASV420	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-1.27	0.01	Female
ASV1341	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Capnocytophaga	-2.01	0.01	Female
ASV241	Actinobacteriota	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces	-1.32	0.01	Female
ASV1091	Patescibacteria	Saccharimonadia	Saccharimonadales	Saccharimonadaceae	TM7x	-1.29	0.01	Female

ANNEXE N

TABLEAU SUPPLÉMENTAIRE 2.3

Tableau supplémentaire 2.3: Analyse discriminante DeSEQ2 d'ASVs bactériens différentiellement abondants dans des communautés bactériennes de larves de TBE entre les tissus intestinaux et les larves entières illustrées à la figure 2.10. Seuls les ASVs représentant plus de 0,01 % de l'ensemble de l'analyse sont présentés dans le tableau et dans la figure 2.10.

	Phylum	Class	Order	Family	Genus	log2FoldChange	% relative abundance	Associated with	Common
ASV1	Actinobacteriota	Actinobacteria	Corynebacteriales	Nocardiaceae	Rhodococcus	-14.01	43.97	gut tissues only	Unique
ASV2	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Enterococcus	8.35	10.04	whole larvae	Common
ASV5	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Mesorhizobium	-11.59	8.19	gut tissues only	Unique
ASV14	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	9.52	2.43	whole larvae	Common
ASV13	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Enterococcus	6.90	1.55	whole larvae	Unique
ASV24	Actinobacteriota	Actinobacteria	Corynebacteriales	Nocardiaceae	Rhodococcus	-8.77	1.16	gut tissues only	Unique
ASV20	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	Endobacter	7.63	0.97	whole larvae	Common
ASV27	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Turicella	-8.49	0.96	gut tissues only	Unique
ASV39	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	8.52	0.92	whole larvae	Common
ASV19	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	8.34	0.90	whole larvae	Common
ASV34	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	Methylocella	8.11	0.90	whole larvae	Common
ASV15	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	7.48	0.64	whole larvae	Common
ASV32	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	8.23	0.61	whole larvae	Unique
ASV16	Firmicutes	Bacilli	Staphylococcales	Staphylococcaceae	Staphylococcus	8.22	0.60	whole larvae	Unique
ASV28	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	Endobacter	5.54	0.55	whole larvae	Common
ASV10	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Lawsonella	-7.66	0.54	gut tissues only	Unique
ASV78	Proteobacteria	Gammaproteobacteria	Burkholderiales	Oxalobacteraceae	Massilia	5.87	0.52	whole larvae	Unique
ASV136	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	Endobacter	2.50	0.46	whole larvae	Common
ASV48	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	3.49	0.39	whole larvae	Common
ASV70	Actinobacteriota	Actinobacteria	Corynebacteriales	Nocardiaceae	Rhodococcus	-7.17	0.38	gut tissues only	Unique
ASV73	Actinobacteriota	Actinobacteria	Corynebacteriales	Nocardiaceae	Rhodococcus	-7.11	0.37	gut tissues only	Unique
ASV58	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	Endobacter	7.22	0.36	whole larvae	Common
ASV115	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	4.04	0.33	whole larvae	Unique
ASV23	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	7.08	0.29	whole larvae	Common
ASV96	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	unassigned	-6.76	0.29	gut tissues only	Unique
ASV135	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	5.45	0.23	whole larvae	Unique
ASV40	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	-6.25	0.20	gut tissues only	Unique
ASV86	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	Methylocella	4.03	0.19	whole larvae	Common
ASV84	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	Methylocella	4.56	0.17	whole larvae	Common
ASV113	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	1.50	0.17	whole larvae	Common
ASV124	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	Alloiococcus	-5.95	0.16	gut tissues only	Unique
ASV61	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-5.95	0.16	gut tissues only	Unique
ASV53	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-5.86	0.15	gut tissues only	Unique
ASV150	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	1.32	0.15	whole larvae	Unique
ASV90	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	2.79	0.14	whole larvae	Common
ASV89	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	3.12	0.13	whole larvae	Unique
ASV168	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	2.14	0.13	whole larvae	Common
ASV25	Proteobacteria	Gammaproteobacteria	Burkholderiales	Burkholderiaceae	Pandoraea	-1.65	0.13	gut tissues only	Common
ASV80	Actinobacteriota	Actinobacteria	Propionibacteriales	Propionibacteriaceae	Cutibacterium	-5.57	0.12	gut tissues only	Unique
ASV248	Acidobacteriota	Acidobacteriae	Acidobacteriales	Acidobacteriaceae (Subgroup 1)	Granulicella	3.16	0.11	whole larvae	Unique
ASV304	Firmicutes	Bacilli	Staphylococcales	Gemellaceae	Gemella	2.94	0.11	whole larvae	Common
ASV166	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	1.78	0.11	whole larvae	Unique
ASV356	Actinobacteriota	Actinobacteria	Micrococcales	Micrococcaceae	Kocuria	2.96	0.10	whole larvae	Common
ASV22	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-5.30	0.10	gut tissues only	Unique
ASV49	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-5.19	0.09	gut tissues only	Unique
ASV74	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-5.09	0.09	gut tissues only	Unique
ASV235	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	unassigned	-4.97	0.08	gut tissues only	Unique
ASV226	Actinobacteriota	Actinobacteria	Corynebacteriales	Nocardiaceae	Rhodococcus	-4.88	0.08	gut tissues only	Unique
ASV263	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	Alloiococcus	-4.87	0.08	gut tissues only	Unique
ASV462	Bacteroidota	Bacteroidia	Flavobacteriales	Weeksellaceae	Chryseobacterium	1.53	0.07	whole larvae	Common
ASV283	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	2.64	0.07	whole larvae	Common
ASV216	Firmicutes	Bacilli	Lactobacillales	Aerococcaceae	Aerococcus	1.72	0.07	whole larvae	Common
ASV44	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	1.29	0.07	whole larvae	Unique
ASV265	Firmicutes	Bacilli	Lactobacillales	Leuconostocaceae	Weissella	-4.63	0.06	gut tissues only	Unique
ASV238	Actinobacteriota	Actinobacteria	Corynebacteriales	Nocardiaceae	Gordonia	-4.56	0.06	gut tissues only	Unique
ASV341	Actinobacteriota	Actinobacteria	Micrococcales	Micrococcaceae	Rothia	-1.87	0.06	gut tissues only	Common
ASV234	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	2.65	0.06	whole larvae	Unique
ASV353	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Enhydrobacter	-1.56	0.06	gut tissues only	Common
ASV312	Actinobacteriota	Actinobacteria	Micrococcales	Micrococcaceae	Kocuria	-4.21	0.05	gut tissues only	Unique
ASV507	Proteobacteria	Gammaproteobacteria	Burkholderiales	Oxalobacteraceae	Massilia	1.77	0.04	whole larvae	Unique
ASV214	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	-4.10	0.04	gut tissues only	Unique
ASV810	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium	-1.67	0.04	gut tissues only	Unique
ASV423	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Turicella	-4.03	0.04	gut tissues only	Unique
ASV394	Actinobacteriota	Actinobacteria	Propionibacteriales	Propionibacteriaceae	Cutibacterium	-3.93	0.04	gut tissues only	Unique
ASV361	Fusobacteriota	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichia	-3.91	0.04	gut tissues only	Unique
ASV91	Firmicutes	Bacilli	Staphylococcales	Staphylococcaceae	Staphylococcus	-3.73	0.03	gut tissues only	Unique

ASV256	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-2.94	0.03	gut tissues only	Unique
ASV492	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Pediococcus	-2.27	0.03	gut tissues only	Unique
ASV1143	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	Acidiphilium	2.26	0.03	whole larvae	Unique
ASV133	Actinobacteriota	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces	-3.54	0.03	gut tissues only	Unique
ASV468	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Capnocytophaga	-3.52	0.03	gut tissues only	Unique
ASV231	Desulfobacterota	Desulfuromonadia	Geoalkalibacteraceae	Geoalkalibacter	unassigned	3.33	0.03	whole larvae	Unique
ASV111	Actinobacteriota	Actinobacteria	Micrococcales	Micrococcaceae	Micrococcus	-3.39	0.03	gut tissues only	Unique
ASV131	Actinobacteriota	Actinobacteria	Propionibacteriales	Propionibacteriaceae	Cutibacterium	-3.39	0.03	gut tissues only	Unique
ASV501	Actinobacteriota	Actinobacteria	Corynebacteriales	Nocardiaceae	Rhodococcus	-3.34	0.02	gut tissues only	Unique
ASV611	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Capnocytophaga	-3.34	0.02	gut tissues only	Unique
ASV450	Deinococcota	Deinococci	Deinococcales	Deinococcaceae	Deinococcus	-1.68	0.02	gut tissues only	Unique
ASV46	Firmicutes	Bacilli	Staphylococcales	Staphylococcaceae	Staphylococcus	-3.24	0.02	gut tissues only	Unique
ASV698	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Turicella	-3.11	0.02	gut tissues only	Unique
ASV1038	Bacteroidota	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas	-1.71	0.02	gut tissues only	Unique
ASV692	Patescibacteria	Saccharimonadia	Saccharimonadales	unassigned	unassigned	-1.78	0.02	gut tissues only	Unique
ASV596	Patescibacteria	Saccharimonadia	Saccharimonadales	Saccharimonadaceae	unassigned	-2.97	0.02	gut tissues only	Unique
ASV52	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	-2.83	0.02	gut tissues only	Unique
ASV742	Actinobacteriota	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces	-2.79	0.02	gut tissues only	Unique
ASV1201	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium	-1.43	0.02	gut tissues only	Unique
ASV880	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Paracoccus	-1.93	0.02	gut tissues only	Unique
ASV365	Proteobacteria	Gammaproteobacteria	Burkholderiales	Neisseriaceae	Neisseria	-2.74	0.02	gut tissues only	Unique
ASV732	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	Bradyrhizobium	-2.74	0.02	gut tissues only	Unique
ASV580	Patescibacteria	Saccharimonadia	Saccharimonadales	Saccharimonadaceae	unassigned	-1.69	0.01	gut tissues only	Unique
ASV1134	Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae	Rhodoluna	-1.72	0.01	gut tissues only	Unique
ASV308	Actinobacteriota	Actinobacteria	Propionibacteriales	Propionibacteriaceae	Cutibacterium	-1.61	0.01	gut tissues only	Unique
ASV201	Actinobacteriota	Actinobacteria	Micrococcales	Dermabacteraceae	Dermabacter	-2.48	0.01	gut tissues only	Unique
ASV1004	Actinobacteriota	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces	-1.77	0.01	gut tissues only	Unique
ASV316	Actinobacteriota	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	-1.51	0.01	gut tissues only	Unique
ASV56	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	-2.33	0.01	gut tissues only	Unique
ASV899	Deinococcota	Deinococci	Deinococcales	Deinococcaceae	Deinococcus	-2.27	0.01	gut tissues only	Unique

ANNEXE O

FIGURE SUPPLÉMENTAIRE 3.1

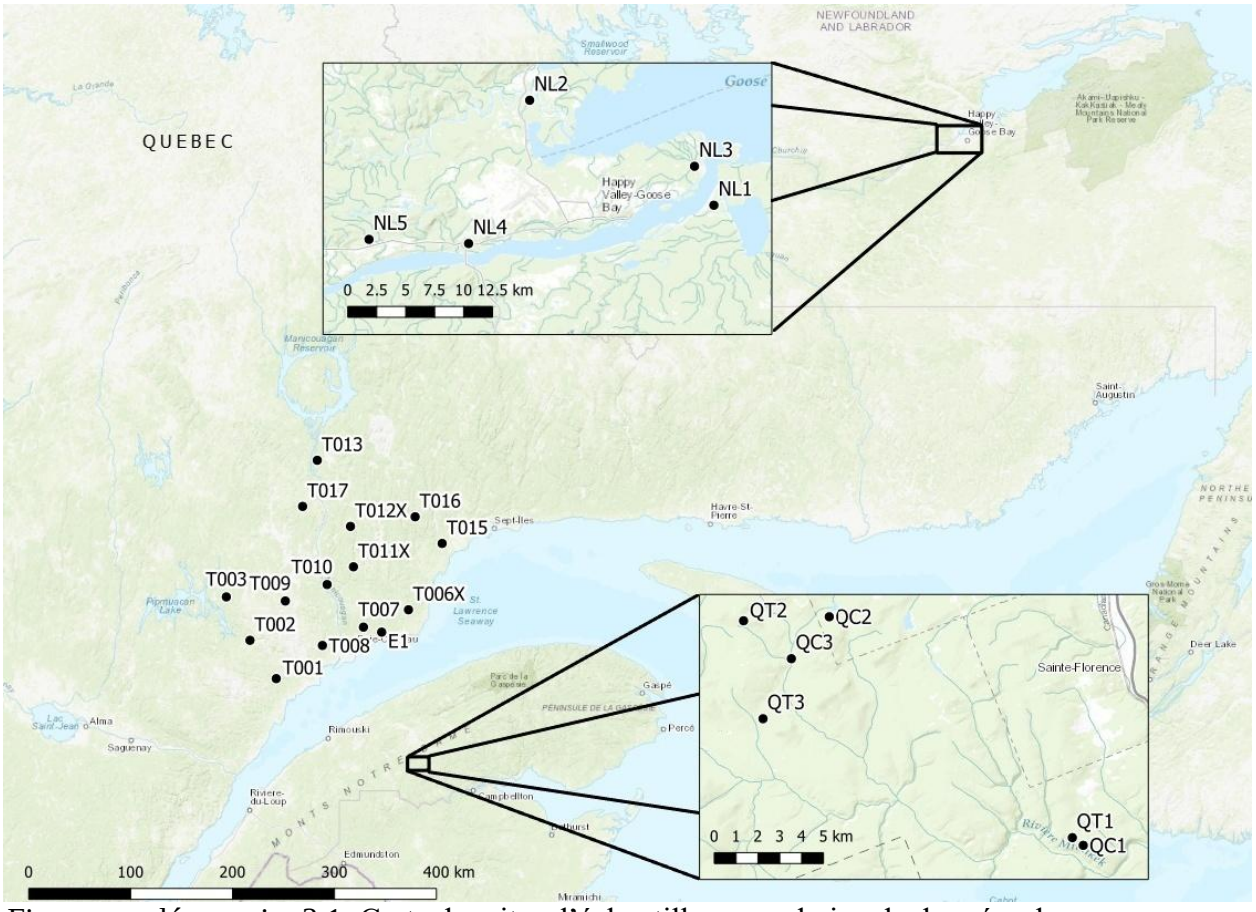

Figure supplémentaire 3.1: Carte des sites d'échantillonnage du jeu de données des larves entières.

ANNEXE P

FIGURE SUPPLÉMENTAIRE 3.2

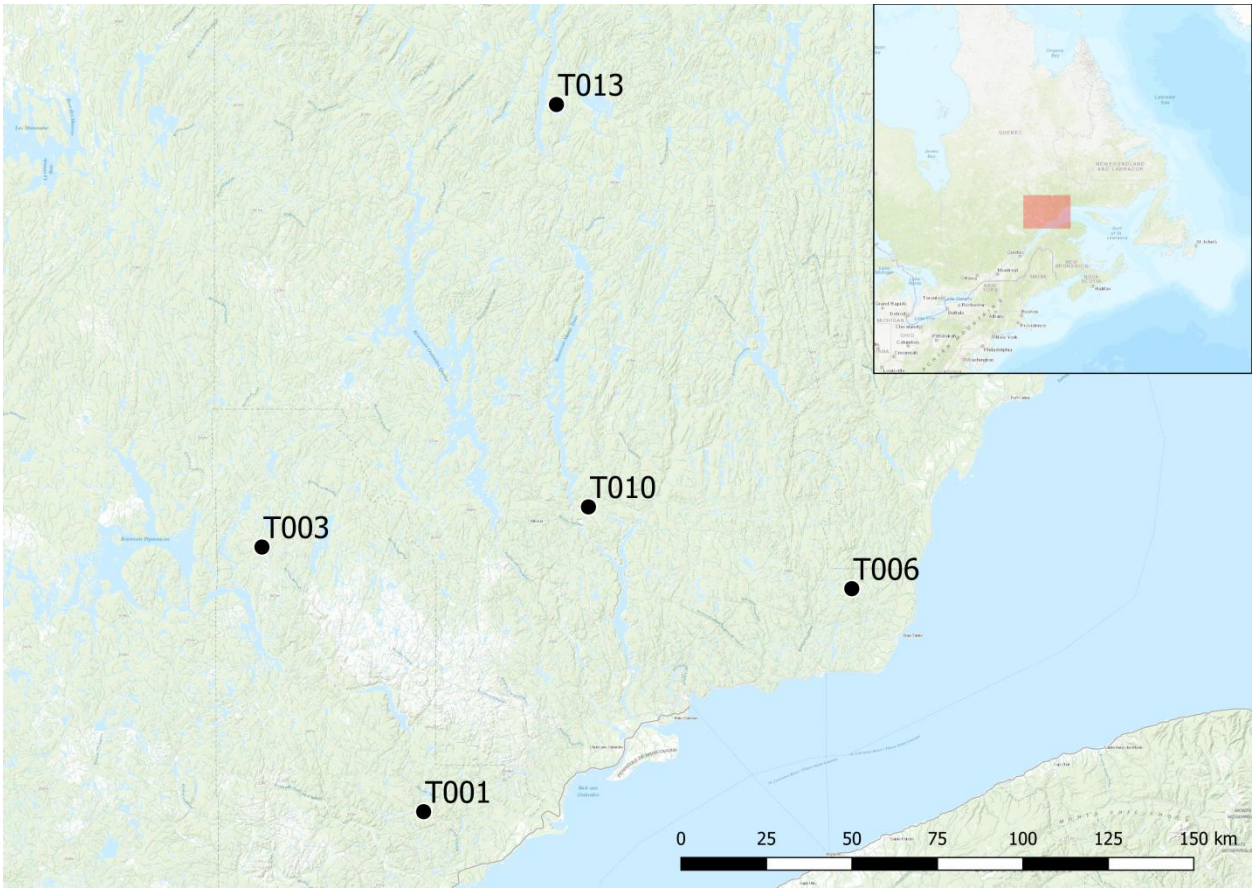

Figure supplémentaire 3.2: Carte des sites d'échantillonnage du jeu de données des tissus intestinaux.

RÉFÉRENCES

Aagaard, K., Petrosino, J., Keitel, W., Watson, M., Katancik, J., Garcia, N., Patel, S., Cutting, M., Madden, T., Hamilton, H., Harris, E., Gevers, D., Simone, G., McInnes, P., & Versalovic, J. (2013). The Human Microbiome Project strategy for comprehensive sampling of the human microbiome and why it matters. *FASEB Journal*, *27*(3), 1012–1022. https://doi.org/10.1096/fj.12-220806

Agler, M. T., Ruhe, J., Kroll, S., Morhenn, C., Kim, S. T., Weigel, D., & Kemen, E. M. (2016). Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *PLoS Biology*, *14*(1), e1002352. https://doi.org/10.1371/journal.pbio.1002352

Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47–97. https://doi.org/10.1103/RevModPhys.74.47

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anand, a A. P., Vennison, S. J., Sankar, S. G., Prabhu, D. I. G., Vasan, P. T., Raghuraman, T., Geoffrey, C. J., & Vendan, S. E. (2010). Isolation and characterization of bacteria from the gut of Bombyx mori that degrade cellulose, xylan, pectin and starch and their impact on digestion. *Journal of Insect Science (Online)*, *10*(107), 107. https://doi.org/10.1673/031.010.10701

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*(1), 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Anderson, M. J. (2006). Distance-Based Tests for Homogeneity of Multivariate Dispersions. *Biometrics*, *62*(1), 245–253. http://www.jstor.org/stable/3695727

Anderson, M. J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA). In *Wiley StatsRef: Statistics Reference Online* (pp. 1–15). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118445112.stat07841

Astudillo-García, C., Bell, J. J., Webster, N. S., Glasl, B., Jompa, J., Montoya, J. M., & Taylor, M. W. (2017). Evaluating the core microbiota in complex communities: A systematic investigation. *Environmental Microbiology*, *19*(4), 1450–1462. https://doi.org/10.1111/1462-2920.13647

Bäckhed, F., Fraser, C. M., Ringel, Y., Sanders, M. E., Sartor, R. B., Sherman, P. M., Versalovic, J., Young, V., & Finlay, B. B. (2012). Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. *Cell Host & Microbe*, *12*(5), 611–622. https://doi.org/10.1016/j.chom.2012.10.012

Baker, F. C., Tsai, L. W., Reuter, C. C., & Schooley, D. A. (1987). In vivo fluctuation of JH, JH acid, and ecdysteroid titer, and JH esterase activity, during development of fifth stadium Manduca sexta. *Insect Biochemistry*, *17*(7), 989–996. https://doi.org/10.1016/0020-1790(87)90108-9

Balser, T. C., Gutknecht, J. L. M., & Liang, C. (2010). How will climate change impact soil microbial communities? In *Soil Microbiology and Sustainable Crop Production* (pp. 373–397). https://doi.org/10.1007/978-90-481-9479-7_10

Baquero, F., & Nombela, C. (2012). The microbiome as a human organ. *Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, *18 Suppl 4*, 2–4. https://doi.org/10.1111/j.1469-0691.2012.03916.x

Basset, Y., Cizek, L., Cuenoud, P., Didham, R. K., Guilhaumon, F., Missa, O., Novotny, V., Odegaard, F., Roslin, T., Schmidl, J., Tishechkin, A. K., Winchester, N. N., Roubik, D. W., Aberlenc, H. P., Bail, J., Barrios, H., Bridle, J. R., Castano-Meneses, G., Corbara, B., … Leponce, M. (2012). Arthropod diversity in a tropical forest. *Science*, *338*(6113), 1481–1484. https://doi.org/10.1126/science.1226727

Beltrán, W., & Wunderle, J. M. (2013). Determinants of tree species preference for foraging by insectivorous birds in a novel Prosopis-Leucaena woodland in Puerto Rico: The role of foliage palatability. *Biodiversity and Conservation*, *22*(9), 2071–2089. https://doi.org/10.1007/s10531-013-0529-x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing Author (s): Yoav Benjamini and Yosef Hochberg Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 57 , No . 1 (1995), Publi. *Journal of the Royal Statistical Society*, *57*(1), 289–300.

Berenbaum, M. (1980). Adaptive Significance of Midgut pH in Larval Lepidoptera. *The American Naturalist*, *115*(1), 29–39.

Best, D. J., & Roberts, D. E. (1975). Algorithm AS 89: The Upper Tail Probabilities of Spearman's Rho. *Applied Statistics*, *24*(3), 377. https://doi.org/10.2307/2347111

Blanchet, F. G., Cazelles, K., & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, *23*(7), 1050–1063. https://doi.org/10.1111/ele.13525

Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*, *92*(5), 1170–1182. https://doi.org/10.1086/228631

Borcard, D., Gillet, F., & Legendre, P. (2011). Numerical Ecology with R. In *Springer*. Springer New York. https://doi.org/10.1007/978-1-4419-7976-6_1

Borcard, D., Legendre, P., & Drapeau, P. (1992). Partialling out the spatial component of ecological variation. *Ecology*, *73*(3), 1045–1055. https://doi.org/10.2307/1940179

Bray, J. R., & Curtis, J. T. (1957). An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs*, *27*(4), 325–349. https://doi.org/10.2307/1942268

Brenner, D. J., Staley, J. T., & Krieg, N. R. (2005). Classification of Procaryotic Organisms and the Concept of Bacterial Speciation. In *Bergey's Manual® of Systematic Bacteriology* (pp. 27–32). https://doi.org/10.1007/0-387-28021-9_4

Breznak, J. a. (1982). Intestinal microbiota of termites and other xylophagous insects. *Annual Review of Microbiology*, *36*, 323–343. https://doi.org/10.1146/annurev.mi.36.100182.001543

Bridgewater, L. C., Zhang, C., Wu, Y., Hu, W., Zhang, Q., Wang, J., Li, S., & Zhao, L. (2017). Gender-based differences in host behavior and gut microbiota composition in response to high fat diet and stress in a mouse model. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-11069-4

Broderick, N. A., Raffa, K. F., Goodman, R. M., & Handelsman, J. (2004). Census of the Bacterial Community of the Gypsy Moth Larval Midgut by Using Culturing and Culture-Independent Methods. *Applied and Environmental Microbiology*, *70*(1), 293–300. https://doi.org/10.1128/AEM.70.1.293-300.2004

Broderick, N. A., Raffa, K., & Handelsman, J. (2006). Midgut bacteria required for Bacillus thuringiensis insecticidal activity. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(41), 15196–15199. https://doi.org/10.1073/pnas.0604865103

Broderick, N. A., Robinson, C. J., McMahon, M. D., Holt, J., Handelsman, J., & Raffa, K. F. (2009). Contributions of gut bacteria to Bacillus thuringiensis-induced mortality vary across a range of Lepidoptera. *BMC Biology*, *7*(1), 11. https://doi.org/10.1186/1741-7007-7-11

Brulc, J. M., Antonopoulos, D. A., Berg Miller, M. E., Wilson, M. K., Yannarell, A. C., Dinsdale, E. A., Edwards, R. E., Frank, E. D., Emerson, J. B., Wacklin, P., Coutinho, P. M., Henrissat, B., Nelson, K. E., & White, B. A. (2009). Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 1948–1953. https://doi.org/10.1073/pnas.0806191105

Buchon, N., Broderick, N. A., & Lemaitre, B. (2013). Gut homeostasis in a microbial world: insights from Drosophila melanogaster. *Nat Rev Microbiol*, *11*(9), 615–626. https://doi.org/10.1038/nrmicro3074

Burleigh, J. S. (1998). *IMPACT OF THE SPRUCE BUDWORM, CHORISTONEURA FUMIFERANA (CLEM.) (LEPIDOPTERA:TORTRICIDAE), ON WHiTE SPRUCE, PICEA GLAUCA (MOENCH) VOSS, IN NORTHEASTE RN BRITISH COLUMBIA*. Simon Fraser University.

Calderón-Cortés, N., Quesada, M., Watanabe, H., Cano-Camacho, H., & Oyama, K. (2011). Endogenous Plant Cell Wall Digestion: A Key Mechanism in Insect Evolution. *Annual Review of Ecology, Evolution, and Systematics*, *43*(1), 120830113150004. https://doi.org/10.1146/annurev-ecolsys-110411-160312

Cáliz, J., Triadó-Margarit, X., Camarero, L., & Casamayor, E. O. (2018). A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(48), 12229–12234. https://doi.org/10.1073/pnas.1812826115

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Cantarel, B. L., Lombard, V., Henrissat, B., Hu, Y., Walker, S., Laine, R., Varki, A., Sharon, N., Varki, A., Hooper, L., Midtvedt, T., Gordon, J., Cantarel, B., Coutinho, P., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B., Henrissat, B., … Sørensen, T. (2012). Complex Carbohydrate Utilization by the Healthy Human Microbiome. *PLoS ONE*, *7*(6), e28742. https://doi.org/10.1371/journal.pone.0028742

Capinera, J. L. (2008). *Encyclopedia of entomology*. Springer.

Cazemier, a E., Hackstein, J. H. P., Op den Camp, H. J. M., Rosenberg, J., & van der Drift, C. (1997). Bacteria in the intestinal tract of differenct species of arthropods. *Microbial Ecology*, *33*, 189–197. https://doi.org/10.1007/s002489900021

Ceja-Navarro, J. A., Vega, F. E., Karaoz, U., Hao, Z., Jenkins, S., Lim, H. C., Kosina, P., Infante, F., Northen, T. R., & Brodie, E. L. (2015). Gut microbiota mediate caffeine detoxification in the primary insect pest of coffee. *Nature Communications*, *6*(May). https://doi.org/10.1038/ncomms8618

Chang, W.-Y., Lantz, V. A., Hennigar, C. R., & MacLean, D. A. (2012). Economic impacts of forest pests: a case study of spruce budworm outbreaks and control in New Brunswick, Canada. *Canadian Journal of Forest Research*, *42*(3), 490–505. https://doi.org/10.1139/x11-190

Chelius, M. K., & Triplett, E. W. (2001). The diversity of archaea and bacteria in association with the roots of Zea mays L. *Microbial Ecology*, *41*(3), 252–263. https://doi.org/10.1007/s002480000087

Chen, B., Teh, B. S., Sun, C., Hu, S., Lu, X., Boland, W., & Shao, Y. (2016). Biodiversity and Activity of the Gut Microbiota across the Life History of the Insect Herbivore Spodoptera littoralis. *Scientific Reports*, *6*(June), 1–14. https://doi.org/10.1038/srep29505

Chen, B., Zhang, N., Xie, S., Zhang, X., He, J., Muhammad, A., Sun, C., Lu, X., & Shao, Y. (2020). Gut bacteria of the silkworm Bombyx mori facilitate host resistance against the toxic effects of organophosphate insecticides. *Environment International*, *143*. https://doi.org/10.1016/j.envint.2020.105886

Cho, I., & Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. *Nature Reviews Genetics*, *13*(4), 260–270. https://doi.org/10.1038/nrg3182

Cohen, H., McFrederick, Q. S., & Philpott, S. M. (2020). Environment Shapes the Microbiome of the Blue Orchard Bee, Osmia lignaria : RRH: Environmental Drivers of Bee Microbiome. *Microbial Ecology*, *80*(4), 897–907. https://doi.org/10.1007/s00248-020-01549-y

Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I., & Knight, R. (2009). Bacterial community variation in human body habitats across space and time. *Science*, *326*(5960), 1694–1697. https://doi.org/10.1126/science.1177486

Crawford, H. S., & Jennings, D. T. (1989). Predation by birds on spruce budworm Choristoneura fumiferana: functional numerical, and total responses. *Ecology*, *70*(1), 152–163. https://doi.org/10.2307/1938422

Csardi, G., & Nepusz, T. (2006). Package “igraph” Title Network Analysis and Visualization. *Interjournal*, 462.

Dalton, T., Dowd, S. E., Wolcott, R. D., Sun, Y., & Watters, C. (2011). An In Vivo Polymicrobial Biofilm Wound Infection Model to Study Interspecies Interactions. *PLoS ONE*, *6*(11), 27317. https://doi.org/10.1371/journal.pone.0027317

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, *6*(1), 221499. https://doi.org/10.1186/s40168-018-0605-2

De Almeida, L. G., De Moraes, L. A. B., Trigo, J. R., Omoto, C., & Cônsoli, F. L. (2017). The gut microbiota of insecticide-resistant insects houses insecticide-degrading bacteria: A potential source for biotechnological exploitation. *PLoS ONE*, *12*(3). https://doi.org/10.1371/journal.pone.0174754

Dekaboruah, E., Suryavanshi, M. V., Chettri, D., & Verma, A. K. (2020). Human microbiome: an academic update on human body site specific surveillance and its possible role. *Archives of Microbiology*, *202*(8), 2147–2167. https://doi.org/10.1007/s00203-020-01931-x

Delmas, E., Besson, M., Brice, M. H., Burkle, L. A., Dalla Riva, G. V., Fortin, M. J., Gravel, D., Guimarães, P. R., Hembry, D. H., Newman, E. A., Olesen, J. M., Pires, M. M., Yeakel, J. D., & Poisot, T. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, *94*(1), 16–36. https://doi.org/10.1111/brv.12433

Dillon, R., & Dillon, V. (2004). The gut bacteria of insects: Nonpathogenic Interactions. *Annual Reviews in Entomology*, *98*, 71–92. https://doi.org/10.1146/annurev.ento.49.061802.123416

Dimopoulos, G. (2003). Insect immunity and its implication in mosquito-malaria interactions. In *Cellular Microbiology* (Vol. 5, Issue 1, pp. 3–14). John Wiley & Sons, Ltd. https://doi.org/10.1046/j.1462-5822.2003.00252.x

Dubilier, N., McFall-Ngai, M., & Zhao, L. (2015). Microbiology: Create a global microbiome effort. *Nature*, *526*(7575), 631–634. https://doi.org/10.1038/526631a

Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002). Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(20), 12917–12922. https://doi.org/10.1073/pnas.192407699

Eisenberg, T., Glaeser, S. P., Kämpfer, P., Schauerte, N., & Geiger, C. (2015). Root sepsis associated with insect-dwelling Sebaldella termitidis in a lesser dwarf lemur (Cheirogaleus medius). *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *108*(6), 1373–1382. https://doi.org/10.1007/s10482-015-0590-4

Engel, P., Martinson, V. G., & Moran, N. a. (2012). Functional diversity within the simple gut microbiota of the honey bee. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1202970109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1202970109

Engel, P., & Moran, N. A. (2013). The gut microbiota of insects - diversity in structure and function. *FEMS Microbiology Reviews*, *37*(5), 699–735. https://doi.org/10.1111/1574-6976.12025

Epstein, H. E., Smith, H. A., Cantin, N. E., Mocellin, V. J. L., Torda, G., & van Oppen, M. J. H. (2019). Temporal variation in the microbiome of acropora coral species does not reflect seasonality. *Frontiers in Microbiology*, *10*(August), 1–14. https://doi.org/10.3389/fmicb.2019.01775

Espenshade, J., Thijs, S., Gawronski, S., Bové, H., Weyens, N., & Vangronsveld, J. (2019). Influence of urbanization on epiphytic bacterial communities of the platanus × hispanica tree leaves in a biennial study. *Frontiers in Microbiology*, *10*(APR), 1–14. https://doi.org/10.3389/fmicb.2019.00675

Eveleigh, E. S., McCann, K. S., McCarthy, P. C., Pollock, S. J., Lucarotti, C. J., Morin, B., McDougall, G. a, Strongman, D. B., Huber, J. T., Umbanhowar, J., & Faria, L. D. B. (2007). Fluctuations in density of an outbreak species drive diversity cascades in food webs. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 16976–16981. https://doi.org/10.1073/pnas.0704301104

Fanson, B. G., Fanson, K. V., & Taylor, P. W. (2013). Sex differences in insect immune function: A consequence of diet choice? *Evolutionary Ecology*, *27*(5), 937–947. https://doi.org/10.1007/s10682-013-9638-y

Fettes, J. P. (1951). *Investigations of sampling techniques for population studies of the spruce budworm on balsam fir in Ontario*. http://www.worldcat.org/title/investigations-of-sampling-techniques-for-population-studies-of-the-spruce-budworm-on-balsam-fir-in-ontario/oclc/62428296

Frankenhuyzen, K. van, Liu, Y., & Tonon, A. (2010). Interactions between Bacillus thuringiensis subsp. kurstaki HD-1 and midgut bacteria in larvae of gypsy moth and spruce budworm. *Journal of Invertebrate Pathology*, *103*(2), 124–131. https://doi.org/10.1016/j.jip.2009.12.008

Franzosa, E. A., Hsu, T., Sirota-Madi, A., Shafquat, A., Abu-Ali, G., Morgan, X. C., & Huttenhower, C. (2015). Sequencing and beyond: Integrating molecular "omics" for microbial community profiling. In *Nature Reviews Microbiology* (Vol. 13, Issue 6, pp. 360–372). Nature Publishing Group. https://doi.org/10.1038/nrmicro3451

Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, *40*(1), 35. https://doi.org/10.2307/3033543

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Friedman, J., & Alm, E. J. (2012). Inferring Correlation Networks from Genomic Survey Data. *PLoS Computational Biology*, *8*(9), e1002687. https://doi.org/10.1371/journal.pcbi.1002687

Fürnkranz, M., Wanek, W., Richter, A., Abell, G., Rasche, F., & Sessitsch, A. (2008). Nitrogen fixation by phyllosphere bacteria associated with higher plants and their colonizing epiphytes of a tropical lowland rainforest of Costa Rica. *The ISME Journal*, *2*(5), 561–570. https://doi.org/10.1038/ismej.2008.14

Germain, M., Kneeshaw, D., De Grandpré, L., Desrochers, M., James, P. M. A., Vepakomma, U., Poulin, J. F., & Villard, M. A. (2021). Insectivorous songbirds as early indicators of future defoliation by spruce budworm. *Landscape Ecology*, *0123456789*. https://doi.org/10.1007/s10980-021-01300-z

Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. In *Ecology Letters* (Vol. 4, Issue 4, pp. 379–391). John Wiley & Sons, Ltd. https://doi.org/10.1046/j.1461-0248.2001.00230.x

Gower, J. C. (1966). Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis. *Biometrika*, *53*(3/4), 325. https://doi.org/10.2307/2333639

Graham, R. I., & Wilson, K. (2012). Male-killing Wolbachia and mitochondrial selective sweep in a migratory African insect. *Bmc Evolutionary Biology*, *12*(1), 204. https://doi.org/Artn 204\rDoi 10.1186/1471-2148-12-204

Grice, E. A., Kong, H. H., Conlan, S., Deming, C. B., Davis, J., Young, A. C., Bouffard, G. G., Blakesley, R. W., Murray, P. R., Green, E. D., Turner, M. L., & Segre, J. A. (2009). Topographical and temporal diversity of the human skin microbiome. *Science*, *324*(5931), 1190–1192. https://doi.org/10.1126/science.1171700

Griffin, J. N., & Silliman, B. R. (2012). Resource Partitioning and Why It Matters Similar Species Compete for Limiting Resources Complete Competitors Cannot Coexist. *Nature Education Knowledge*, *3*(10), 49. http://cmore.soest.hawaii.edu/summercourse/2013/documents/Gifford_06-07/Griffin_et_al_2012_Resource_Partitioning_and_Why_It_Matters .pdf

Halsch, C. A., Shapiro, A. M., Fordyce, J. A., Nice, C. C., Thorne, J. H., Waetjen, D. P., & Forister, M. L. (2021). Insects and recent climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(2), 1–9. https://doi.org/10.1073/PNAS.2002543117

Hammer, T. J., Janzen, D. H., Hallwachs, W., Jaffe, S. P., & Fierer, N. (2017). Caterpillars lack a resident gut microbiome. *Proceedings of the National Academy of Sciences*, *114*(36), 9641–9646. https://doi.org/10.1073/pnas.1707186114

Hammer, T. J., McMillan, W. O., & Fierer, N. (2014). Metamorphosis of a butterfly-associated bacterial community. *PLoS ONE*, *9*(1), e86995. https://doi.org/10.1371/journal.pone.0086995

Hammer, T. J., Sanders, J. G., & Fierer, N. (2019). Not all animals need a microbiome. In *FEMS Microbiology Letters* (Vol. 366, Issue 10, p. 117). Oxford Academic. https://doi.org/10.1093/femsle/fnz117

Han, G., Lee, H. J., Jeong, S. E., Jeon, C. O., & Hyun, S. (2017). Comparative Analysis of Drosophila melanogaster Gut Microbiota with Respect to Host Strain, Sex, and Age. *Microbial Ecology*, *74*(1), 207–216. https://doi.org/10.1007/s00248-016-0925-3

Harmon, L. J., & Glor, R. E. (2010). Poor statistical performance of the mantel test in phylogenetic comparative analyses. *Evolution*, *64*(7), 2173–2178. https://doi.org/10.1111/j.1558-5646.2010.00973.x

Harrison, J. F. (2001). Insect acid-base physiology. *Annual Review of Entomology*.

Hennigar, C. R., Maclean, D. A., Quiring, D. T., & Kershaw, J. A. (2008). *Differences in Spruce Budworm Defoliation among Balsam Fir and Black Spruce*. *54*(2), 158–166.

Hernández-Gómez, O. (2020). Climate change disturbs wildlife microbiomes. *Nature Climate Change 2020 10:11*, *10*(11), 981–982. https://doi.org/10.1038/s41558-020-00923-5

Hernandez, D. J., David, A. S., Menges, E. S., Searcy, C. A., & Afkhami, M. E. (2021). Environmental stress destabilizes microbial networks. *ISME Journal*, *15*(6), 1722–1734. https://doi.org/10.1038/s41396-020-00882-x

Huang, J. H., Jing, X., & Douglas, A. E. (2015). The multi-tasking gut epithelium of insects. *Insect Biochemistry and Molecular Biology*, *67*, 15–20. https://doi.org/10.1016/j.ibmb.2015.05.004

Human Microbiome Project Consortium, T. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, *486*. https://doi.org/10.1038/nature11234

Hung, K.-L. J., Kingston, J. M., Albrecht, M., Holway, D. A., & Kohn, J. R. (2018). The worldwide importance of honey bees as pollinators in natural habitats. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1870), 20172140. https://doi.org/10.1098/rspb.2017.2140

James, P. M. A., Robert, L.-E., Wotton, B. M., Martell, D. L., & Fleming, R. A. (2017). Lagged cumulative spruce budworm defoliation affects the risk of fire ignition in Ontario, Canada. *Ecological Applications*. https://doi.org/10.1002/eap.1463

Jardon, Y., Morin, H., & Dutilleul, P. (2003). *Périodicité et synchronisme des épidémies de la tordeuse des bourgeons de l'épinette au Québec*. https://doi.org/10.1139/X03-108

Jašarević, E., Morrison, K. E., & Bale, T. L. (2016). Sex differences in the gut microbiome - Brain axis across the lifespan. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 371, Issue 1688). Royal Society of London. https://doi.org/10.1098/rstb.2015.0122

Kennedy, S. R., Tsau, S., Gillespie, R., & Krehenwinkel, H. (2020). Are you what you eat? A highly transient and prey-influenced gut microbiome in the grey house spider Badumna longinqua. *Molecular Ecology*, *29*(5), 1001–1015. https://doi.org/10.1111/mec.15370

Khalifa, A. Y. Z., Alsyeeh, A. M., Almalki, M. A., & Saleh, F. A. (2016). Characterization of the plant growth promoting bacterium, Enterobacter cloacae MSR1, isolated from roots of non-nodulating Medicago sativa. *Saudi Journal of Biological Sciences*, *23*(1), 79–86. https://doi.org/10.1016/j.sjbs.2015.06.008

Kikuchi, Y., Hayatsu, M., Hosokawa, T., Nagayama, A., Tago, K., & Fukatsu, T. (2012). Symbiont-mediated insecticide resistance. *Proceedings of the National Academy of Sciences*, *109*(22), 8618–8622. https://doi.org/Doi 10.1073/Pnas.1200231109

Kishore, D., Birzu, G., Hu, Z., Delisi, C., Korolev, K. S., & Segrè, D. (2020). Inferring microbial co-occurrence networks from amplicon data: a systematic evaluation. *BioRxiv*, 2020.09.23.309781. https://doi.org/10.1101/2020.09.23.309781

Kostic, A. D. A. D., Xavier, R. J. R. J., Gevers, D., Sartor, R. B., Swidsinski, A., Ladhoff, A., Pernthaler, A., al., et, Loftus, E. V., Jostins, L., Ripke, S., Weersma, R. K., al., et, Khor, B., Gardet, A., Xavier, R. J. R. J., Molodecky, N. A., Soon, I. S., Rabi, D. M., … Tillinghast, G. W. (2014). The microbiome in inflammatory bowel disease: Current status and the future ahead. *Gastroenterology*, *146*(6), 1489–1499. https://doi.org/10.1053/j.gastro.2014.02.009

Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., & Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, *29*(5), 592–599. https://doi.org/10.1111/1365-2435.12345

Kumar, M., Kour, D., Yadav, A. N., Saxena, R., Rai, P. K., Jyoti, A., & Tomar, R. S. (2019). Biodiversity of methylotrophic microbial communities and their potential role in mitigation of abiotic stresses in plants. In *Biologia* (Vol. 74, Issue 3, pp. 287–308). De Gruyter. https://doi.org/10.2478/s11756-019-00190-6

Kurtz, Z. D., Müller, C. L., Miraldi, E. R., Littman, D. R., Blaser, M. J., & Bonneau, R. A. (2015). Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLoS Computational Biology*, *11*(5), e1004226. https://doi.org/10.1371/journal.pcbi.1004226

Laforest-Lapointe, I., Messier, C., & Kembel, S. W. (2016). Host species identity, site and time drive temperate tree phyllosphere bacterial community structure. *Microbiome*, *4*(1), 27. https://doi.org/10.1186/s40168-016-0174-1

Landry, M., Comeau, A. M., Derome, N., Cusson, M., & Levesque, R. C. (2015). Composition of the Spruce Budworm (Choristoneura fumiferana) Midgut Microbiota as Affected by Rearing Conditions. *PloS One*, *10*(12), e0144077. https://doi.org/10.1371/journal.pone.0144077

Landry, M., James, P. M. A., Kneeshaw, D., & Kembel, S. W. (2022). Spruce budworm bacterial communities vary among sites and host tree species in a boreal landscape. *Journal of Biogeography*. https://doi.org/10.1111/JBI.14299

Lankau, E. W., Hong, P. Y., & MacKie, R. I. (2012). Ecological drift and local exposures drive enteric bacterial community differences within species of Galápagos iguanas. *Molecular Ecology*, *21*(7), 1779–1788. https://doi.org/10.1111/j.1365-294X.2012.05502.x

Larroque, J., Johns, R., Canape, J., Morin, B., & James, P. M. A. (2020). *Spatial genetic structure at the leading edge of a spruce budworm outbreak: The role of dispersal in outbreak spread*. https://doi.org/10.1016/j.foreco.2020.117965

Layeghifard, M., Hwang, D. M., & Guttman, D. S. (2017). Disentangling Interactions in the Microbiome: A Network Perspective. *Trends in Microbiology*, *25*(3), 217–228. https://doi.org/10.1016/j.tim.2016.11.008

Lederberg, J., & McCray, A. (2001). 'Ome Sweet 'Omics-- A Genealogical Treasury of Words. *Scientist*.

Legendre, P., Borcard, D., & Peres-Neto, P. R. (2005). Analyzing beta diversity: Partitioning the spatial variation of community composition data. *Ecological Monographs*, *75*(4), 435–450. https://doi.org/10.1890/05-0549

Legendre, P., & Fortin, M. J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, *10*(5), 831–844. https://doi.org/10.1111/j.1755-0998.2010.02866.x

Legendre, P., Fortin, M. J., & Borcard, D. (2015). Should the Mantel test be used in spatial analysis? *Methods in Ecology and Evolution*, *6*(11), 1239–1247. https://doi.org/10.1111/2041-210X.12425

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*(2), 271–280. https://doi.org/10.1007/s004420100716

Legrand, F., Chen, W., Cobo-Díaz, J. F., Picot, A., & Le Floch, G. (2019). Co-occurrence analysis reveal that biotic and abiotic factors influence soil fungistasis against Fusarium graminearum. *FEMS Microbiology Ecology*, *95*(5). https://doi.org/10.1093/femsec/fiz056

Levin, D., Raab, N., Pinto, Y., Rothschild, D., Zanir, G., Godneva, A., Mellul, N., Futorian, D., Gal, D., Leviatan, S., Zeevi, D., Bachelet, I., & Segal, E. (2021). Diversity and functional landscapes in the microbiota of animals in the wild. *Science*, *372*(6539), eabb5352. https://doi.org/10.1126/science.abb5352

Li, L., Liu, Z., Zhang, M., Meng, D., Liu, X., Wang, P., Li, X., Jiang, Z., Zhong, S., Jiang, C., & Yin, H. (2020). Insights into the Metabolism and Evolution of the Genus Acidiphilium, a Typical Acidophile in Acid Mine Drainage. *MSystems*, *5*(6). https://doi.org/10.1128/MSYSTEMS.00867-20

Li, S. peng, Wang, P., Chen, Y., Wilson, M. C., Yang, X., Ma, C., Lu, J., Chen, X. yong, Wu, J., Shu, W. sheng, & Jiang, L. (2020). Island biogeography of soil bacteria and fungi: similar patterns, but different mechanisms. *ISME Journal*, *14*(7), 1886–1896. https://doi.org/10.1038/s41396-020-0657-8

Lin, H., & Peddada, S. Das. (2020). Analysis of compositions of microbiomes with bias correction. *Nature Communications*, *11*(1), 1–11. https://doi.org/10.1038/s41467-020-17041-7

Lisboa, F. J. G., Peres-Neto, P. R., Chaer, G. M., Jesus, E. da C., Mitchell, R. J., Chapman, S. J., & Berbara, R. L. L. (2014). Much beyond Mantel: Bringing Procrustes Association Metric to the Plant and Soil Ecologist's Toolbox. *PLoS ONE*, *9*(6), e101238. https://doi.org/10.1371/journal.pone.0101238

Logares, R., Deutschmann, I. M., Junger, P. C., Giner, C. R., Krabberød, A. K., Schmidt, T. S. B., Rubinat-Ripoll, L., Mestre, M., Salazar, G., Ruiz-González, C., Sebastián, M., De Vargas, C., Acinas, S. G., Duarte, C. M., Gasol, J. M., & Massana, R. (2020). Disentangling the mechanisms shaping the surface ocean microbiota. *Microbiome*, *8*(1), 1–17. https://doi.org/10.1186/s40168-020-00827-8

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

MacLean, D., Beaton, K., Porter, K., MacKinnon, W., & Budd, M. (2002). Potential wood supply losses to spruce budworm in New Brunswick estimated using the Spruce Budworm Decision Support System. In *Forestry Chronicle* (Vol. 78, Issue 5, pp. 739–750).

MacLean, D., & Piene, H. (1995). Spatial and temporal patterns of balsam fir mortality in spaced and unspaced stands caused by spruce budworm defoliation. *Canadian Journal of Forest Research*, *25*(6), 902–911. https://doi.org/10.1139/x95-099

Magan, N., & McLeod, A. R. (1991). *Effects of Atmospheric Pollutants on Phyllosphere Microbial Communities* (pp. 379–400). Springer, New York, NY. https://doi.org/10.1007/978-1-4612-3168-4_19

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis* (p. 521). Academic Press.

Marrec, R., Pontbriand-Pare, O., Legault, S., & James, P. M. A. (2018). Spatiotemporal variation in drivers of parasitoid metacommunity structure in continuous forest landscapes. *Ecosphere*, *9*(1). https://doi.org/10.1002/ecs2.2075

Martineau, R. (1985). Tordeuse des bourgeons de l'épinette. *Insectes Nuisibles Des Forêts de l'est Du Canada*, 283.

Martinson, J. N. V., Pinkham, N. V., Peters, G. W., Cho, H., Heng, J., Rauch, M., Broadaway, S. C., & Walk, S. T. (2019). Rethinking gut microbiome residency and the Enterobacteriaceae in healthy human adults. *ISME Journal*, *13*(9), 2306–2318. https://doi.org/10.1038/s41396-019-0435-7

Martiny, A. C., Treseder, K., & Pusch, G. (2013). Phylogenetic conservatism of functional traits in microorganisms. *ISME Journal*, *7*(4), 830–838. https://doi.org/10.1038/ismej.2012.160

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Øvreås, L., Reysenbach, A. L., Smith, V. H., & Staley, J. T. (2006). Microbial biogeography: Putting microorganisms on the map. In *Nature Reviews Microbiology* (Vol. 4, Issue 2, pp. 102–112). https://doi.org/10.1038/nrmicro1341

Martiny, J. B. H., Jones, S. E., Lennon, J. T., & Martiny, A. C. (2015). Microbiomes in light of traits: A phylogenetic perspective. *Science*, *350*(6261). https://doi.org/10.1126/science.aac9323

Matarrita-Carranza, B., Moreira-Soto, R. D., Murillo-Cruz, C., Mora, M., Currie, C. R., & Pinto-Tomas, A. A. (2017). Evidence for widespread associations between neotropical hymenopteran insects and Actinobacteria. *Frontiers in Microbiology*, *8*(OCT), 1–17. https://doi.org/10.3389/fmicb.2017.02016

McCullagh, P., & Nelder, J. A. (2019). Generalized Linear Models. In *Regression Analysis with Application G.B. Wetherill* (Issue 2). Routledge. https://doi.org/10.1201/9780203753736

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, *298*(5594), 824–827. https://doi.org/10.1126/science.298.5594.824

Ministère des Forêts, de la F. et des P. (2020). (2020). *Aires infestées par la tordeuse des bourgeons de l'épinette au Québec en 2020*. http://nfdp.ccfm.org/fr/data/insects.php

Mohajeri, M. H., Brummer, R. J. M., Rastall, R. A., Weersma, R. K., Harmsen, H. J. M., Faas, M., & Eggersdorfer, M. (2018). The role of the microbiome for human health: from basic science to clinical applications. *European Journal of Nutrition*, *57*(1), 1–14. https://doi.org/10.1007/s00394-018-1703-4

Moran, N. A., Ochman, H., & Hammer, T. J. (2019). Evolutionary and Ecological Consequences of Gut Microbial Communities. *Annual Review of Ecology, Evolution, and Systematics*, *50*, 451–475. https://doi.org/10.1146/annurev-ecolsys-110617-062453

Murray, J. L., Connell, J. L., Stacy, A., Turner, K. H., & Whiteley, M. (2014). Mechanisms of Synergy in Polymicrobial Infections. *Journal of Microbiology*, *52*(3), 188–199. https://doi.org/10.1007/s12275-014-4067-3

Nakashima, Y., & Hirose, Y. (2003). Sex differences in foraging behaviour and oviposition site preference in an insect predator, Orius sauteri. *Entomologia Experimentalis et Applicata*, *106*(2), 79–86. https://www.researchgate.net/profile/Yoshitaka_Nakashima/publication/227702972_Sex_differences_in_foraging_behaviour_and_oviposition_site_preference_in_an_insect_predator_Orius_sauteri/links/5a5fe026458515b4377b85fa/Sex-differences-in-foraging-behaviour-a

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P., & Ferrenberg, S. (2013). Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews*, *77*(3), 342–356. https://doi.org/10.1128/MMBR.00051-12

NIH Human Microbiome Portfolio Analysis Team. (2019). *NIH Human Microbiome Portfolio Analysis Team*.* 1–19. https://doi.org/10.1186/s40168-019-0620-y

Ojha, A., & Zhang, W. (2019). A comparative study of microbial community and dynamics of Asaia in the brown planthopper from susceptible and resistant rice varieties. *BMC Microbiology*, *19*(1), 139. https://doi.org/10.1186/s12866-019-1512-9

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R. B., Simpson, G., Solymos, P., Stevens, M., & Wagner, H. (2016). *vegan: Community Ecology R package.* https://cran.r-project.org/package=vegan

Otti, O. (2015). Genitalia-associated microbes in insects. In *Insect Science* (Vol. 22, Issue 3, pp. 325–339). Blackwell Publishing Ltd. https://doi.org/10.1111/1744-7917.12183

Palma, L., Muñoz, D., Berry, C., Murillo, J., & Caballero, P. (2014). Bacillus thuringiensis Toxins: An Overview of Their Biocidal Activity. *Toxins*, *6*, 3296–3325. https://doi.org/10.3390/toxins6123296

Pang, A. S. D. (2010). Possible Role of Plant Allelochemical in Clearance of Bacteria from the Gut of Spruce Budworm, Choristoneura fumiferana. *The Open Microbiology Journal*, *4*, 26–29. https://doi.org/10.2174/1874285801004010026

Panteli, N., Mastoraki, M., Lazarina, M., Chatzifotis, S., Mente, E., Kormas, K. A., & Antonopoulou, E. (2021). Configuration of gut microbiota structure and potential functionality in two teleosts under the influence of dietary insect meals. *Microorganisms*, *9*(4). https://doi.org/10.3390/microorganisms9040699

Pantoja, A., Smith, C. M., & Robinson, J. F. (1986). Effects of the Fall Armyworm (Lepidoptera: Noctuidae) on Rice Yields. *Journal of Economic Entomology*, *79*(5), 1324–1329. https://doi.org/10.1093/jee/79.5.1324

Pattabhiramaiah, M., Reddy, M. S., & Brueckner, D. (2012). *Presence of Wolbachia in the Lepidopteran insect pests of mulberry*. *1*(1), 16–22. https://www.researchgate.net/profile/Mahesh_Pattabhiramaiah/publication/284430864_Presence_of_Wolbachia_in_the_Lepidopteran_insect_pests_of_mulberry/links/5653000708ae4988a7af3608.pdf?origin=publication_detail

Pereira, S. I. A., Pires, C., Henriques, I., Correia, A., Magan, N., & Castro, P. M. L. (2015). Assessment of rhizospheric culturable bacteria of Phragmites australis and Juncus effusus from polluted sites. *Journal of Basic Microbiology*, *55*(10), 1179–1190. https://doi.org/10.1002/jobm.201500010

Peres-Neto, P. R., & Jackson, D. A. (2001). How well do multivariate data sets match? The advantages of a procrustean superimposition approach over the Mantel test. *Oecologia*, *129*(2), 169–178. https://doi.org/10.1007/s004420100720

Peterson, B. F., & Scharf, M. E. (2016). Lower Termite Associations with Microbes: Synergy, Protection, and Interplay. *Frontiers in Microbiology*, *7*, 422. https://doi.org/10.3389/fmicb.2016.00422

Phalnikar, K., Kunte, K., & Agashe, D. (2019). Disrupting butterfly caterpillar microbiomes does not impact their survival and development. *Proceedings of the Royal Society B*, *286*(1917). https://doi.org/10.1098/RSPB.2019.2438

Piper, F. I., Gundale, M. J., & Fajardo, A. (2015). Extreme defoliation reduces tree growth but not C and N storage in a winter-deciduous species. *Annals of Botany*, *115*(7), 1093–1103. https://doi.org/10.1093/aob/mcv038

Pollock, J., Glendinning, L., Wisedchanwet, T., & Watson, M. (2018). The madness of microbiome: Attempting to find consensus "best practice" for 16S microbiome studies. In *Applied and Environmental Microbiology* (Vol. 84, Issue 7). https://doi.org/10.1128/AEM.02627-17

Pothier, D., & Mailly, D. (2006). Stand-level prediction of balsam fir mortality in relation to spruce budworm defoliation. *Canadian Journal of Forest Research*, *36*(7), 1631–1640. https://doi.org/10.1139/x06-062

Pureswaran, D. S., Johns, R., Heard, S. B., & Quiring, D. (2016). Paradigms in eastern spruce budworm (Lepidoptera: Tortricidae) population ecology: A century of debate. *Environmental Entomology*, *45*(6), 1333–1342. https://doi.org/10.1093/ee/nvw103

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

R Development Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.r-project.org/

Raulo, A., Allen, B. E., Troitsky, T., Husby, A., Firth, J. A., Coulson, T., & Knowles, S. C. L. (2021). Social networks strongly predict the gut microbiota of wild mice. *ISME Journal*. https://doi.org/10.1038/s41396-021-00949-3

Régnière, J., St-Amant, R., & Duval, P. (2012). Predicting insect distributions under climate change from physiological responses: Spruce budworm as an example. *Biological Invasions*, *14*(8), 1571–1586. https://doi.org/10.1007/s10530-010-9918-1

Rossmassler, K., Dietrich, C., Thompson, C., Mikaelyan, A., Nonoh, J. O., Scheffrahn, R. H., Sillam-Dussès, D., & Brune, A. (2015). Metagenomic analysis of the microbiota in the highly compartmented hindguts of six wood- or soil-feeding higher termites. *Microbiome*, *3*(1), 56. https://doi.org/10.1186/s40168-015-0118-1

Rousk, J., & Bengtson, P. (2014). *The Microbial Regulation of Global Biogeochemical Cycles*. https://doi.org/10.3389/978-2-88919-297-7

Royama, T. (1984). Population Dynamics of the Spruce Budworm Choristoneura Fumiferana. *Ecological Monographs*, *54*(4), 429–462. https://doi.org/10.2307/1942595

Russell, D. A., Radwan, S. M., Irving, N. S., Jones, K. A., & Downham, M. C. A. (1993). Experimental assessment of the impact of defoliation by Spodoptera littoralis on the growth and yield of Giza ’75 cotton. *Crop Protection*, *12*(4), 303–309. https://doi.org/10.1016/0261-2194(93)90051-J

Russell, J. A., Funaro, C. F., Giraldo, Y. M., Goldman-Huertas, B., Suh, D., Kronauer, D. J. C., Moreau, C. S., & Pierce, N. E. (2012). A Veritable Menagerie of Heritable Bacteria from Ants, Butterflies, and Beyond: Broad Molecular Surveys and a Systematic Review. *PLoS ONE*, *7*(12). https://doi.org/10.1371/journal.pone.0051027

Schäfer, A., Konrad, R., Kuhnigk, T., Kämpfer, P., Hertel, H., & König, H. (1996). Hemicellulose-degrading bacteria and yeasts from the termite gut. *The Journal of Applied Bacteriology*, *80*(5), 471–478. https://doi.org/10.1111/j.1365-2672.1996.tb03245.x

Schwarz, M. T., Kneeshaw, D., & Kembel, S. W. (2018). The Gut-Associated Microbiome of the Eastern Spruce Budworm Does Not Influence Larval Growth or Survival. *BioRxiv*, 330928. https://doi.org/10.1101/330928

Scully, E. D., Geib, S. M., Hoover, K., Tien, M., Tringe, S. G., Barry, K. W., Glavina del Rio, T., Chovatia, M., Herr, J. R., & Carlson, J. E. (2013). Metagenomic Profiling Reveals Lignocellulose Degrading System in a Microbial Community Associated with a Wood-Feeding Beetle. *PLoS ONE*, *8*(9), 1–22. https://doi.org/10.1371/journal.pone.0073827

Segata, N., Baldini, F., Pompon, J., Garrett, W. S., Truong, D. T., Dabiré, R. K., Diabaté, A., Levashina, E. A., & Catteruccia, F. (2016). The reproductive tracts of two malaria vectors are populated by a core microbiome and by gender-and swarm-enriched microbial biomarkers. *Scientific Reports*, *6*. https://doi.org/10.1038/srep24207

Sepulveda, J., & Moeller, A. H. (2020). The Effects of Temperature on Animal Gut Microbiomes. *Frontiers in Microbiology*, *11*(March), 1–9. https://doi.org/10.3389/fmicb.2020.00384

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Simmons, B. I., Cirtwill, A. R., Baker, N. J., Wauchope, H. S., Dicks, L. V., Stouffer, D. B., & Sutherland, W. J. (2019). Motifs in bipartite ecological networks: uncovering indirect interactions. *Oikos*, *128*(2), 154–170. https://doi.org/10.1111/oik.05670

Six, D. L. (2013). The Bark Beetle Holobiont: Why Microbes Matter. *Journal of Chemical Ecology*, *39*(7), 989–1002. https://doi.org/10.1007/s10886-013-0318-8

Smith, A. H., Łukasik, P., O'Connor, M. P., Lee, A., Mayo, G., Drott, M. T., Doll, S., Tuttle, R., Disciullo, R. A., Messina, A., Oliver, K. M., & Russell, J. A. (2015). Patterns, causes and consequences of defensive microbiome dynamics across multiple scales. *Molecular Ecology*, *24*(5), 1135–1149. https://doi.org/10.1111/mec.13095

Smith, C. C., Srygley, R. B., Healy, F., Swaminath, K., & Mueller, U. G. (2017). Spatial structure of the mormon cricket gut microbiome and its predicted contribution to nutrition and immune function. *Frontiers in Microbiology*, *8*(MAY), 801. https://doi.org/10.3389/fmicb.2017.00801

Stehlikova, Z., Kostovcikova, K., Kverka, M., Rossmann, P., Dvorak, J., Novosadova, I., Kostovcik, M., Coufal, S., Srutkova, D., Prochazkova, P., Hudcovic, T., Kozakova, H., Stepankova, R., Rob, F., Juzlova, K., Hercogova, J., Tlaskalova-Hogenova, H., & Zakostelska, Z. J. (2019). Crucial role of microbiota in experimental psoriasis revealed by a gnotobiotic mouse model. *Frontiers in Microbiology*, *10*(FEB), 236. https://doi.org/10.3389/fmicb.2019.00236

Stocks, B. J. (1987). Fire Potential in the Spruce Budworm-damaged Forests of Ontario. *The Forestry Chronicle*, *63*(1), 8–14. https://doi.org/10.5558/tfc63008-1

Stothart, M. R., Greuel, R. J., Gavriliuc, S., Henry, A., Wilson, A. J., McLoughlin, P. D., & Poissant, J. (2021). Bacterial dispersal and drift drive microbiome diversity patterns within a population of feral hindgut fermenters. *Molecular Ecology*, *30*(2), 555–571. https://doi.org/10.1111/mec.15747

Thermo Fisher. (2010). *RNAlater Preserves Bacterial Gene Expression Profiles for Array Analysis - JP*. https://www.thermofisher.com/ca/en/home/references/ambion-tech-support/rna-isolation/tech-notes/rnalater-preserves-bacterial-gene-expression-profiles-for-array-.html

Tu, Q., Yan, Q., Deng, Y., Michaletz, S. T., Buzzard, V., Weiser, M. D., Waide, R., Ning, D., Wu, L., He, Z., & Zhou, J. (2020). Biogeographic patterns of microbial co-occurrence ecological networks in six American forests. *Soil Biology and Biochemistry*, *148*(June), 107897. https://doi.org/10.1016/j.soilbio.2020.107897

Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., & Gordon, J. I. (2007). The Human Microbiome Project. In *Nature* (Vol. 449, Issue 7164, pp. 804–810). Nature Publishing Group. https://doi.org/10.1038/nature06244

Valeri, F., & Endres, K. (2021). How biological sex of the host shapes its gut microbiota. In *Frontiers in Neuroendocrinology* (Vol. 61, p. 100912). Academic Press Inc. https://doi.org/10.1016/j.yfrne.2021.100912

van der Heijden, M. G. A., & Hartmann, M. (2016). Networking in the Plant Microbiome. In *PLoS Biology* (Vol. 14, Issue 2, p. e1002378). Public Library of Science. https://doi.org/10.1371/journal.pbio.1002378

Vellend, M. (2016). *The Theory of Ecological Communities - Full Book*. https://press.princeton.edu/titles/10914.html

Vemuri, R., Sylvia, K. E., Klein, S. L., Forster, S. C., Plebanski, M., Eri, R., & Flanagan, K. L. (2019). The microgenderome revealed: sex differences in bidirectional interactions between the microbiota, hormones, immunity and disease susceptibility. In *Seminars in Immunopathology* (Vol. 41, Issue 2, pp. 265–275). https://doi.org/10.1007/s00281-018-0716-7

Vilanova, C., Baixeras, J., Latorre, A., & Porcar, M. (2016). The generalist inside the specialist: Gut bacterial communities of two insect species feeding on toxic plants are dominated by Enterococcus sp. *Frontiers in Microbiology*, *7*(JUN). https://doi.org/10.3389/fmicb.2016.01005

Voigt, A. Y., Costea, P. I., Kultima, J. R., Li, S. S., Zeller, G., Sunagawa, S., & Bork, P. (2015). Temporal and technical variability of human gut metagenomes. *Genome Biology*, *16*(1). https://doi.org/10.1186/s13059-015-0639-8

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). *Locally Downscaled and Spatially Customizable Climate Data for Historical and Future Periods for North America*. https://doi.org/10.1371/journal.pone.0156720

Warnecke, F., Luginbühl, P., Ivanova, N., Ghassemian, M., Richardson, T. H., Stege, J. T., Cayouette, M., McHardy, A. C., Djordjevic, G., Aboushadi, N., Sorek, R., Tringe, S. G., Podar, M., Martin, H. G., Kunin, V., Dalevi, D., Madejska, J., Kirton, E., Platt, D., … Leadbetter, J. R. (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, *450*(November), 560–565. https://doi.org/10.1038/nature06269

Willey, J. M. (2014). *Prescott's microbiology* (p. 51).

Willink, B., & Svensson, E. I. (2017). Intra-and intersexual differences in parasite resistance and female fitness tolerance in a polymorphic insect. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1847). https://doi.org/10.1098/rspb.2016.2407

Wright, D. E. (1986). Damage and loss in yield of wheat crops caused by the australian plague locust, chortoicetes terminifera (Walker). *Australian Journal of Experimental Agriculture*, *26*(5), 613–618. https://doi.org/10.1071/EA9860613

Xia, X., Zheng, D., Zhong, H., Qin, B., Gurr, G. M., Vasseur, L., Lin, H., Bai, J., He, W., & You, M. (2013). DNA Sequencing Reveals the Midgut Microbiota of Diamondback Moth, Plutella xylostella (L.) and a Possible Relationship with Insecticide Resistance. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0068852

Yan, B., Li, J., Xiao, N., Qi, Y., Fu, G., Liu, G., & Qiao, M. (2016). Urban-development-induced Changes in the Diversity and Composition of the Soil Bacterial Community in Beijing. *Scientific Reports*, *6*(December). https://doi.org/10.1038/srep38811

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), 1–11. https://doi.org/10.1038/s41598-018-34127-x

Zhao, X., Yang, Y., Feng, K., Wang, X., Liu, B., Xie, G., & Xing, D. (2021). Self-regulating microbiome networks ensure functional resilience of biofilms in sand biofilters during manganese load fluctuations. *Water Research*, *188*, 116473. https://doi.org/10.1016/j.watres.2020.116473

Zhu, L., Wu, Q., Dai, J., Zhang, S., & Wei, F. (2011). Evidence of cellulose metabolism by the giant panda gut microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(43), 17714–17719. https://doi.org/10.1073/pnas.1017956108