

Le microbiote bactérien cuticulaire des fourmis de Guyane : pouvoir antibiotique et écologie des communautés

Caroline Birer

► To cite this version:

Caroline Birer. Le microbiote bactérien cuticulaire des fourmis de Guyane : pouvoir antibiotique et écologie des communautés. Chimie organique. Université de Guyane, 2017. Français. <NNT : 2017YANE0003>. <tel-01537793>

HAL Id: tel-01537793

<https://tel.archives-ouvertes.fr/tel-01537793>

Submitted on 13 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Thèse

En vue de l'obtention du grade de

Docteur de l'université de Guyane

Discipline : *Chimie*

Spécialité : *Chimie et biologie des produits naturels*



Présentée et soutenue publiquement le 6 avril 2017 par :

CAROLINE BIRER

Le microbiote bactérien cuticulaire des fourmis de Guyane : pouvoir antibiotique et écologie des communautés

JURY

Jérôme CHAVE	HDR, DR	Président du Jury, Rapporteur
Jean-Luc WOLFENDER	HDR, PU	Rapporteur
Heidy SCHIMANN	CR	Examineur
Grégory GENTA-JOUE	HDR, MCF	Examineur
Christophe DUPLAIS	HDR, CR	Directeur de Thèse
Bruno FIGADERE	HDR, DR	Co-Directeur de Thèse
Lucie ZINGER	MCF	Co-encadrant
Niklas TYSKLIND	CR	Co-encadrant

Remerciements

Ce manuscrit de thèse et l'ensemble du travail de recherche présentés ici ne sont pas les fruits des efforts d'une seule personne, mais les efforts conjoints de plusieurs personnes autour de moi qui m'ont formés et m'ont apportés leurs aides et leurs soutiens.

J'aimerais en premier lieu remercier les membres du jury de cette thèse qui m'ont fait l'honneur de bien vouloir juger ces travaux. Veuillez trouver ici l'expression de toute ma gratitude :

Dr. Jérôme Chave
Pr. Jean-Luc Wolfender
Dr. Grégory Genta-Jouve
Dr. Heidy Schimann

Je remercie grandement mon directeur de thèse, le Dr. Christophe Duplais. Je suis extrêmement reconnaissante à Christophe Duplais pour la formation à la recherche, par la recherche et pour la recherche, qu'il m'a prodiguée, pour son encadrement de tous les jours, pour ses encouragements pour que je me forme auprès d'autres collègues. Quel coach moral tu as fait, merci pour ton soutien sans faille ! Que de discussions dans ce bureau partagé, ta passion pour la recherche est sans limite et a été le catalyseur de mon ouverture d'esprit scientifique. Je retiendrai toujours ton « Persévérance, Humilité et Créativité » de mon premier jour de thèse.

Je remercie aussi pleinement le Dr. Bruno Figadère qui a co-dirigé cette thèse et m'a accueilli à trois reprises dans son laboratoire (UMR BioCIS) où j'ai été encadrée par le Dr. Alexandre Maciuk, à qui j'adresse également mes remerciements.

Je remercie énormément le Dr. Lucie Zinger, qui au fil du temps et de la thèse est devenue une co-encadrante de choc et a beaucoup contribué à mon apprentissage de la recherche. Tu as toujours été très réactive pour répondre à mes questions et tes réponses nettes et riches en explications ont finalement permis que je me forme à distance en écologie. Je te suis très reconnaissante du temps précieux que tu as consacré à ma formation en métabarcoding, à manipuler R, à la conception de mon premier article et à m'inculquer toutes ces notions passionnantes en écologie des communautés.

Je remercie également beaucoup mon co-encadrant, le Dr. Niklas Tysklind qui m'a très chaleureusement accueilli dans son laboratoire de génétique à EcoFoG où j'ai eu pendant quelques mois mon deuxième bureau. Merci pour ton soutien lorsque les échecs PCR remplissaient mon cahier de laboratoire, pour tes nombreux conseils et pour ton aide précieuse pendant la semaine de terrain aux Nouragues.

Je remercie amplement le Dr. Grégory Genta-Jouve, parce que tu as aussi été un co-encadrant à mon sens et que tu as aussi pris part à ma thèse au fil du temps. Je te remercie pour ta franchise, la très bonne formation en métabolomique et en statistiques que tu m'as prodiguée. Je te remercie pour m'avoir accueillie de nombreuses fois à la faculté de pharmacie de l'Université Paris Descartes où j'avais mon troisième bureau.

Ce projet de thèse m'a permis de profiter d'une mobilité extrême, ce résultat reste à confirmer mais il semblerait que l'interdisciplinarité rime avec mobilité. Aussi, je remercie toutes les équipes des laboratoires qui m'ont chaleureusement accueillies en me laissant bénéficier de leurs installations et pour les nombreuses discussions qui s'y sont déroulées. En Guyane, je remercie le laboratoire Interactions Virus-Hôtes à l'Institut Pasteur de la Guyane, le laboratoire de Génétique Ecologique et celui de Biologie des Interactions à EcoFoG à Kourou, le laboratoire Hydreco à Petit Saut et le laboratoire des Nouragues. Je remercie très chaleureusement Valérie et Eliane qui sont les rouages du laboratoire de génétique de Kourou, je vous remercie pour votre accueil, votre patience et pour la très bonne formation en biologie moléculaire que vous m'avez donné. De l'autre côté de l'Atlantique, je remercie le laboratoire BioCIS à la faculté de pharmacie de Châtenay-Malabry, notamment Karine et ses conseils sur la spectrométrie de masse, et le laboratoire C-TAC à la faculté de pharmacie de l'Université Paris-Descartes. Je remercie également Josie Santini, notre gestionnaire d'EcoFoG, grâce à qui tous les bons de commande ont été envoyés et tous les billets d'avion ont été réservés.

Cette forte mobilité qui m'a ravie durant toute ma thèse s'est faite à partir du Laboratoire des Substances Naturelles Amazoniennes d'EcoFoG basé sur le campus de l'Institut Pasteur à côté de la pointe Buzaré à Cayenne. C'est dans ce laboratoire que j'ai passé la majorité de ma thèse, à couper les climatiseurs pour le bon développement des actinobactéries, à prendre le 4x4 pour aller sur le terrain ou pour « monter à Kourou », à manipuler dans le petit laboratoire presque P2 et surtout à bénéficier de l'aide et des conseils de tous ses occupants. Les Dr. Yannick Estevez, Emeline Houël et Nadine Amusant, je vous remercie beaucoup pour votre soutien et Yannick je te remercie énormément pour ton ingéniosité à réparer avec brio les machines sur lesquelles le sort s'acharne, tu m'as sauvé plus d'une fois. Je remercie également les techniciens et post-docs présents lors de ma première année de thèse Mickael, Arnaud, Isabelle, Jean et Louis-Jérôme : merci pour les conseils de grand-frère/sœur que vous m'avez prodigués sur la Guyane et sur ma thèse. Je remercie beaucoup Claudiane pour son soutien et je te souhaite beaucoup de courage pour finir ta thèse sans me faire trop de souci pour toi ! Je remercie Cyrielle qui est arrivée un peu tard pour que l'on manipule ensemble, mais qui je suis sûre est bien à sa place dans ce laboratoire. Et oui Damien je t'intègre au paragraphe du labo,

tu y es bien non ? Damien, je te remercie pour ton soutien, pour tes réactions toujours très pertinentes sur ma thèse qui ont beaucoup contribué à ce travail, pour t'assurer que ça va bien tous les jours, pour tes conseils en microbiologie, en biologie moléculaire et pour m'avoir formée à la phylogénie.

Je remercie tous les doctorants d'EcoFoG à Kourou qui contribuent à la sympathie qui règne sur le campus, et tous les chercheurs d'EcoFoG qui savent être disponibles pour les étudiants en répondant aux nombreuses questions. Je remercie spécialement Axel Touchard qui m'a appris tout ce que je sais sur les fourmis tropicales. Je te remercie de m'avoir accompagnée de si nombreuses fois sur le terrain ! Je remercie Audrey Sagne, pour nos discussions lors des aller-retours entre Cayenne et Kourou et pour ton aide sur R, et des coups de main pour quelques manip aussi !

Je remercie également les stagiaires, Marie-Claire, Marie-Hélène, Sacha et Franciskar qui ont travaillé avec moi dans la joie et la bonne humeur et avec une motivation sans faille. Trois d'entre eux sont issus de la licence « BBB » pour Biologie, Biochimie et Biotechnologie de l'Université de Guyane, j'ai été ravie de vous transmettre mes connaissances lors de sessions de terrain inoubliables et pluvieuses, au laboratoire lors des innombrables bains que nous avons donnés aux fourmis et derrière l'écran lors des rédactions de rapport de stage.

L'écriture de ce manuscrit s'est terminée par de nombreuses relectures. En plus des très bons conseils de mes encadrants et de leurs corrections, je tiens à remercier Yannick, Damien, Stan, Yanouk et J-B pour avoir donné leur avis et traqué les coquilles. Vos relectures m'ont été très précieuses, je vous remercie beaucoup.

Je remercie mes anciens professeurs et mes anciens encadrants de stage, de m'avoir initiée à la recherche. Je remercie particulièrement Anne Mandeau et Georges Massiot qui m'ont convaincue que cette thèse en Guyane à EcoFoG avec Christophe Duplais était une belle opportunité.

Je remercie Corrie Moreau pour l'intérêt qu'elle porte à mon travail, pour nos discussions très intéressantes aux Nouragues, pour ses encouragements, ses nombreux conseils et sa relecture précieuse de l'article.

Je remercie mes collocs, mes voisins, ma famille, mes amis et amies en Guyane et en métropole pour leur soutien au quotidien. Cette thèse en Guyane est une expérience très riche tant sur le plan personnel que professionnel, je suis très reconnaissante à toutes ces personnes qui ont contribué à mon épanouissement pendant cette formation à la recherche.

Titre de la thèse : Le microbiote bactérien cuticulaire des fourmis de Guyane : pouvoir antibiotique et écologie des communautés

Résumé : Le microbiote bactérien cuticulaire des fourmis (Hymenoptera : Formicidae) est connu pour avoir un rôle défensif chez ces insectes sociaux, notamment chez les fourmis attines (Formicidae : Attini) grâce l'utilisation de molécules antimicrobiennes produites par des actinobactéries cuticulaires. Dans le cadre de cette thèse, nous avons étudié le microbiote bactérien des fourmis de Guyane en utilisant différentes approches en chimie des produits naturels et en écologie moléculaire. Le premier chapitre décrit l'isolement, l'identification, la culture et l'évaluation biologique de 43 actinobactéries cuticulaires de fourmis de Guyane. Les tests d'antagonismes des souches isolées et l'activité antibiotique des extraits de culture contre des micro-organismes pathogènes humains sont présentés ainsi que l'identification d'un dipeptide cyclique (Cyclo(LPro-LPhe)) antimicrobien qui a été isolé à partir d'une souche proche de *Streptomyces thioluteus*. Par ailleurs, la mise en œuvre de réseaux moléculaires appliqués à une analyse par UPLC/MS/MS de cocultures d'actinobactéries a permis d'explorer la diversité des métabolites produits dans ces conditions. Le deuxième chapitre présente une étude méthodologique pour comparer quatre méthodes d'extraction d'ADN, en termes de richesse et de composition du microbiote bactérien cuticulaire, par séquençage haut débit à partir des espèces *Atta cephalotes* et *Pseudomyrmex penetrator*. Les résultats du métabarcoding ADN mettent en lumière deux méthodes d'extraction et révèlent des différences inter- et intraspécifiques dans la composition des communautés bactériennes cuticulaires. Enfin, le chapitre trois décrit la composition du microbiote bactérien cuticulaire des espèces *Camponotus femoratus* et *Crematogaster levior* dans les jardins de fourmis. Les résultats soulignent l'acquisition d'une partie du microbiote dans l'environnement. En parallèle l'analyse métabolomique des cuticules montre à contrario une plus grande spécificité liée à l'espèce de fourmi. Les recherches futures axées sur les stratégies d'analyses statistiques combinant le métabarcoding et la métabolomique sont discutées.

Mots clefs : fourmis, microbiote, cuticule, actinobactéries, cocultures, cyclodipeptides, métabolomique, réseaux moléculaires, métabarcoding, extraction d'ADN, jardins de fourmis.

Title: Bacterial microbiota of ant's cuticle in French Guiana: antibiotic activities and community ecology

Summary: The bacterial microbiota of ants (Hymenoptera: Formicidae) is known to have a defensive role in social insects, particularly for leaf-cutting ants (Formicidae: Attini) due to the use of antimicrobial molecules produced by cuticular actinobacteria. In this thesis, we studied the bacterial microbiota of ants in French Guyana using different approaches based on natural products chemistry and molecular ecology. The first chapter describes the isolation, identification, culture and biological evaluations of 43 cuticular actinobacteria. Antagonism bioassays of isolated strains and antibiotic activities of the culture extracts against human pathogens are presented as well as the identification of an antimicrobial cyclic dipeptide (Cyclo (LPro-LPhe)) isolated from a strain close to *Streptomyces thioluteus*. Moreover, the implementation of molecular networks applied to UPLC/MS/MS analysis of actinobacterial cocultures allowed us to explore the diversity of metabolites produced under these conditions. The second chapter presents a methodological study to evaluate the capacity of four DNA extraction methods, in terms of richness and composition of the cuticular bacterial microbiota, in high-throughput sequencing from *Atta cephalotes* and *Pseudomyrmex penetrator*. The results of metabarcoding highlight two methods of extraction and reveal inter- and intraspecific differences in the composition of cuticular bacterial communities. Finally, chapter three describes the composition of the cuticular bacterial microbiota of *Camponotus femoratus* and *Crematogaster levior* in ant garden and the results reveal the acquisition in the environment of a part of the microbiota. In parallel, metabolomic analyses of ant's cuticle show, on the contrary, a greater specificity related to the ant species. Future researches focusing on statistical analysis strategies combining metabarcoding and metabolomics data are discussed.

Keywords: ants, microbiota, cuticle, actinobacteria, cocultures, cyclodipeptides, metabolomics, molecular networks, metabarcoding, DNA extraction, ant garden.

Table des figures

Figure 1 : Schéma du cycle de vie des actinobactéries incluant un schéma d'une coupe transversale d'actinobactérie avec des hyphes vivants (Barka et al. 2016).	34
Figure 2 : Phylogénie moléculaire de 97 séquences du génome du phylum Actinobacteria (Barka et al. 2016).	35
Figure 3 : Exemples de molécules actives issues d'actinobactéries (Barka et al. 2016).	37
Figure 4 : Streptomyces griseus et la molécule de streptomycine.....	38
Figure 5 : Structure chimique de molécules isolées à partir de la culture d'actinobactéries collectées sur la cuticule d'insectes.	40
Figure 6 : Position de la glande métapleurale (Yek & Mueller 2011).	41
Figure 7 : Mutualisme des attines (Ramadhar et al. 2014).....	42
Figure 8 : Molécules actives isolées à partir d'actinobactéries de la surface de fourmis attines.	43
Figure 9 : Ouverture de la glande métapleurale chez Paraponera clavata (Yek & Mueller 2011).....	47
Figure 10 : Structure de la valinomycine isolée du Streptomyces associé à Acromyrmex echiniator.....	48
Figure 11 : Images microscopiques (A) et images MALDI (B) d'Acromyrmex echiniator. Des spots de valinomycine (rouge) sont déposés à côté de la fourmi et permettent d'estimer la concentration (Schoenian et al. 2011).	48
Figure 12 : Photographie de Camponotus femoratus (à droite) et de Crematogaster levior (à gauche) (auteur : Alain Dejean).	49
Figure 13 : Deux stades de développement des jardins de fourmis avec Camponotus femoratus et Crematogaster levior, mature à gauche et jeune à droite avec les épiphytes Codonantho calcarata et Aechmea mertensii (Orivel & Leroy 2011).	50
Figure 14 : Structure chimique de la crématoénone (Menzel et al. 2013).	51
Figure 15 : Schéma du spectromètre de masse SYNAPT G2 HDMS (Waters Corporation, Manchester, UK).....	57
Figure 16 : Représentation tridimensionnelle d'un chromatogramme obtenu par Liquid Chromatography-Mass Spectrometry.	58
Figure 17 : Amplification en phase solide (HTS ; Illumina).....	63
Figure 18 : L'intégration de données « multi-omics » pour faciliter l'analyse intégrée d'un microbiote (Beale et al. 2016).	65

Figure 19 : 1) entrée du nid de <i>Paraponera clavata</i> (auteur : Sacha Cataldo), 2) prélèvement sur le terrain (auteur : Christophe Duplais) et 3) mise en culture sur boîtes avec le laboratoire mobile (auteur : Christophe Duplais).....	68
Figure 20 : Structure chimique de la géosmine.	70
Figure 21 : Photos de boîtes de Pétri avec 1) plusieurs souches bactériennes, 2) la souche isolée CBCu566 et 3) CBCu788.	71
Figure 22 : Schéma bilan de la méthode d'extraction des métabolites.	74
Figure 23 : Formule de calcul du pourcentage d'inhibition de la croissance bactérienne, DO drug = mesure d'absorbance du puits contenant la suspension bactérienne et l'extrait testé, DO milieu = mesure d'absorbance du milieu de culture stérile, DO parasite seul = mesure d'absorbance de la suspension bactérienne seule.....	75
Figure 24 : Chromatogramme HPLC de l'extrait brut de la souche 497, indications des pics présents dans les fractions actives F4, F5 et F6.....	86
Figure 25 : Structure moléculaire de la dicétopipérazine (R, S)Pro-(R, S)Phe isolée..	87
Figure 26 : Corrélation COSY et HMBC de la dicétopipérazine (R, S)Pro-(R, S)Phe isolée.....	87
Figure 27 : Comparaison des spectres de dichroïsme circulaire expérimental et modélisé.	88
Figure 28 : Exemple de structure chimique de dicétopipérazines et de leurs activités biologiques respectives.	89
Figure 29 : Réseau moléculaire et structure chimique des 2,5 dicétopipérazines identifiées dans la souche 497.....	91
Figure 30 : Biosynthèse de l'auréothine (Adermann et al. 2004).....	92
Figure 31 : Spectre expérimental MS ² de l'auréothine. (Auréothine standard : MS ¹ :398.15 [M+H] ⁺ , MS ² : 370,1 262,0 203,0) (Shao et al. 2011).....	92
Figure 32 : Composition des six assemblages de cocultures, avec les codes des six souches d'actinobactéries au centre.	94
Figure 33 : Chromatogrammes HPLC UV superposés de la CoC4 et des deux monocultures 581 et 5513.	96
Figure 34 : Réseaux moléculaires des six souches en monoculture (gris) et de la coculture CoC1 (rose).	97
Figure 35 : Réseaux moléculaires des trois souches en monoculture (gris) et de la coculture CoC2 (bleu).	98
Figure 36 : Réseaux moléculaires des six souches en monoculture (gris) et des six cocultures CoC1, CoC2, CoC3, CoC4, CoC5, CoC6 (vert).....	99

Figure 37 : Structures chimique des Antimycine A1 à A4.....	99
Figure 38 : Photos des fourmis <i>Pseudomyrmex penetrator</i> (à gauche) (auteur : Caroline Birer) <i>Atta cephalotes</i> (à droite) (auteur : Alex Wild http://www.alexanderwild.com).	101
Figure 39 : Experimental design for evaluating bacterial cuticular microbiota using different DNA extraction protocols. The design was applied to two ant species are <i>A. cephalotes</i> and <i>P. penetrator</i> , and the extraction protocols tested are QIAamp: Qiagen QIAamp DNA mini kit, Dneasy: Qiagen DNeasy blood and tissue kit, PC: Phenol-chloroform, MO BIO: MO BIO UltraClean tissues and cells.....	106
Figure 40 : Effect of DNA extraction protocols on the number of reads (a), OTUs (b) and diversity coverage (c-d). In (a-b), number of reads and OTUs are represented for <i>A. cephalotes</i> (blue), <i>P. penetrator</i> (yellow) ant species with cleaned data and a logarithmic scale (y axis). The figure includes DNA extraction controls (grey) for comparative purposes. (c-d) represent the rarefaction curves on averaged bacterial communities for each DNA extraction methods for (a) <i>A. cephalotes</i> and (b) <i>P. penetrator</i>	112
Figure 41 : Taxonomic composition of bacterial cuticular microbiota and effect of DNA extraction methods on bacterial beta diversity. (a) Taxonomic composition of bacterial cuticular microbiota of <i>A. cephalotes</i> (blue) and <i>P. penetrator</i> (yellow). Only bacterial clades with a total relative abundance greater than 0.1% are shown. (b) Similarities of samples in their community composition according to DNA extraction methods. Dot size represents the sequencing depth of each sample. Only samples with more than 100 reads are included.....	114
Figure 42 : Effect of the number of individuals and individual body size on bacterial beta diversity and taxonomic composition of bacterial cuticular microbiota. (a) Bacterial community composition similarities between <i>A. cephalotes</i> and <i>P. penetrator</i> using QIAamp DNA extraction protocol. (b) Taxonomic composition of bacterial cuticular microbiota of <i>A. cephalotes</i> major workers (brown) and minor workers (orange). Only bacterial clades with a total relative abundance greater than 0.1% are shown.	116
Figure 43 : Carte de Guyane présentant les trois sites de collecte, https://www.geoportail.gouv.fr/mentions-legales	124
Figure 44 : Photographies illustrant les étapes du prélèvement et du lavage, (auteur : Christophe Duplais).....	125
Figure 45 : Représentation graphique de l'Analyse en Composante principale (ACP) de toute la matrice de données de métabolomique.	127
Figure 46 : Histogramme des valeurs d'intensité des ions de tous les QC de la matrice de données de métabolomique.	128

Figure 47 : Courbe de raréfaction (a) et richesse (nombre d'OTUs) à partir de données raréfiées (b) pour <i>Camponotus femoratus</i> , <i>Crematogaster levior</i> et les nids prélevés sur trois sites de prélèvement : à la montagne de Kaw (violet), à Petit Saut (rose) et à la station scientifique des Nouragues (vert).....	133
Figure 48 : Observation des données brutes de métabolomique représenté par le nombre d'ions pour <i>Camponotus femoratus</i> et <i>Crematogaster levior</i> sur trois sites de prélèvement : à la montagne de Kaw (violet), à Petit Saut (rose) et à la station scientifique des Nouragues (vert).	134
Figure 49 : Diagramme de Venn des OTUs pour les deux espèces de fourmis <i>Camponotus femoratus</i> et <i>Crematogaster levior</i> ainsi que leur nid à Kaw (a), à Petit Saut (b) et aux Nouragues (c).	135
Figure 50 : Similarités de composition des communautés d'OTUs pour tous les échantillons de nids. La taille des points représente la profondeur de séquençage (nombre de reads) pour chaque échantillon.....	137
Figure 51 : Similarités de composition des communautés d'OTUs pour tous les individus. La taille des points représente la profondeur de séquençage (nombre de reads) pour chaque individu.	138
Figure 52 : sPLS-DA avec la variable catégorielle espèce.....	139
Figure 53 : Clustered image map, représentation graphique des 10 OTUs sélectionnés par le composant 1 du modèle sPLS-DA selon la variable catégorielle espèce (<i>Camponotus femoratus</i> en vert et <i>Crematogaster levior</i> en violet).	141
Figure 54 : sPLS-DA avec la variable catégorielle site de prélèvement.	142
Figure 55 : Contribution plot des OTUs sélectionnés par le composant 1 (a)) et par le composant 2 (b)) de la sPLS-DA selon la variable catégorielle du site de prélèvement (Kaw en bleu, Petit Saut en rose et les Nouragues en vert).	143
Figure 56 : Similarités de composition des communautés d'ions pour les individus des deux espèces <i>Camponotus femoratus</i> et <i>Crematogaster levior</i>	144
Figure 57 : sPLS-DA avec la variable catégorielle espèce.....	145
Figure 58 : Contribution plot des 50 premiers ions sélectionnés par le composant 1 de la sPLS-DA selon la variable catégorielle espèce (<i>Camponotus femoratus</i> en vert et <i>Crematogaster levior</i> en violet).	146
Figure 59 : Analyse de masse de Kendricks des 50 premiers ions sélectionnés par le composant 1 de la sPLS-DA selon l'espèce. Les ions les plus abondants pour une espèce sont de la couleur de l'espèce (<i>Camponotus femoratus</i> en vert et <i>Crematogaster levior</i> en violet)	147

Figure 60 : Schéma de la méthode cross-streak.	158
Figure 61 : Schéma d'une plaque de 96 puits de test antimicrobien.....	161

Table des tableaux

Table 1 : L'occurrence des classes de CHCs (en pourcentage) dans 4 groupes d'Hyménoptères. Le nombre d'espèces considérées dans chaque groupe (N) est indiqué (van Wilgenburg, Symonds, et Elgar 2011).	46
Table 2 : Description des méthodes d'évaluation de l'activité antibiotique.	53
Table 3 : Description des souches pathogènes choisies.	73
Table 4 : Liste des espèces de fourmis (habitat, lieux et dates de collectes) et identification des 43 souches d'actinobactéries isolées (souche la plus proche et pourcentage de similarité selon EzTaxon).	79
Table 5 : Pourcentage d'inhibition des 43 souches d'actinobactéries contre 6 souches pathogènes, inhibition totale = +++ (rouge), inhibition forte entre 50% et 100% = + (orange) et inhibition faible entre 15% et 50% = +/- (jaune) et absence d'inhibition = - (blanc)	82
Table 6 : Pourcentage d'inhibition des extraits purifiés des 43 souches d'actinobactéries contre quatre souches tests pathogènes, inhibition totale = +++ (rouge), inhibition forte entre 50% et 100% = + (orange) et inhibition faible entre 15% et 50% = +/- (jaune) et absence d'inhibition = - (blanc). Photos d'actinobactérie après 5 semaines, de haut en bas : CBCu555, CBCu497 et CBCu782.	85
Table 7 : Résultats des CI_{50} des fractions F4, F5 et F6.	86
Table 8 : Table des résultats RMN de la dicétopipérazine (R, S)Pro-(R, S)Phe isolée.	88
Table 9 : Pourcentage d'inhibition des extraits purifiés des 6 souches d'actinobactéries et des 6 cocultures contre 4 souches tests pathogènes, inhibition totale = +++ (rouge), inhibition forte entre 50% et 100% = + (orange) et inhibition faible entre 15% et 50% = +/- (jaune) et absence d'inhibition = - (blanc).	95
Table 10 : Liste des 10 OTUs sélectionnés par le composant 1 du modèle sPLS-DA selon la variable catégorielle espèce classés par ordre d'importance décroissant.	141
Table 11 : Informations chimiques de la base de données Metlin pour 5 ions de <i>Camponotus femoratus</i>	148
Table 12 : Caractéristiques des 9 espèces de fourmis pour l'étude à plus grande échelle.	154

Sommaire

Remerciements	3
Table des figures.....	9
Table des tableaux	14
Sommaire.....	16
Avant-propos	22
Introduction générale.....	24
Etat de l'art général des modèles d'étude.....	28
I. Ecologie des communautés microbiennes.....	28
A. Généralités	28
1) Communautés microbiennes, « espèce » et fonctions	28
2) Distribution et règles d'assemblage.....	29
B. Les interactions hôtes-micro-organismes : Le microbiote.....	30
1) Définitions	30
2) Acquisition et transmission	31
3) Exemples de systèmes hôte-microbiote.....	32
II. Les actinobactéries	33
A. Présentation.....	33
B. Intérêts pharmaceutiques	36
C. Mutualisme entre actinobactéries cuticulaires et leurs hôtes.....	38
III. Le microbiote bactérien de la cuticule des fourmis.....	40
A. Contexte	40
B. Les attines	41
C. Autres fourmis	43
D. Assemblages de ces microbiotes bénéfiques	44
IV. Les métabolites de la cuticule des fourmis.....	45
A. Les hydrocarbures cuticulaires	45

B.	Les glandes exocrines : cas de la glande métapleurale	46
C.	Les métabolites des micro-organismes cuticulaires.....	47
V.	Les jardins de fourmis : un cas de parabiosis	49
	Etat de l'art général des méthodes.....	52
I.	Méthodes d'évaluation de l'activité antibiotique	52
A.	Présentation et contexte	52
B.	La méthode cross-streak	54
II.	Techniques de chimie analytique	54
A.	La chromatographie	54
B.	La spectrométrie de masse	55
C.	Apports et principe des réseaux moléculaires.....	58
D.	La métabolomique	59
1)	Définitions	59
2)	Prétraitement des données	60
3)	Contrôle de l'acquisition avec les QC	60
III.	Technique de génétique	61
A.	Métabarcoding	61
B.	Description de la technique de séquençage HTS	62
C.	Les approches multi-omics :	64
	Chapitre 1. Pouvoir antibiotique d'actinobactéries cuticulaires de fourmis.....	66
I.	Introduction	66
II.	Partie Expérimentale.....	67
A.	Collectes des fourmis	67
B.	Isolement des actinobactéries	68
1)	Lavage	68
2)	Mise en culture	69
3)	Isolement des actinobactéries :	70
C.	Identification par séquençage de l'ARNr 16S	72

D.	Pouvoir antibiotique des souches d'actinobactéries isolées :	72
1)	Méthode cross-streak.....	73
2)	Méthode broth dilution	74
(a)	Préparation des extraits.....	74
(b)	Pourcentage d'inhibition des extraits	75
E.	Isolement et identification d'une molécule chimique.....	75
1)	Souche 497 : culture liquide, extraction, isolement.....	75
2)	Détermination structurale : RMN, Spectrométrie de masse et Dichroïsme circulaire	76
F.	Les cocultures	77
G.	Analyse de spectrométrie de masse : réseau moléculaire	77
III.	Résultats et discussions	78
A.	Souches d'actinobactéries isolées	78
B.	Pouvoir antibiotique des actinobactéries isolées :	79
1)	Antagonismes aux quatre bactéries pathogènes :	79
2)	Antagonismes à <i>Candida albicans</i> et <i>Beauveria bassiana</i> :	80
C.	Activité antibiotique des métabolites extraits des souches d'actinobactéries	83
D.	Cyclodipeptides issus de la souche 497 : identification structurale et réseau moléculaire	86
1)	Déréplication	86
2)	Identification structurale.....	87
3)	Les dicétopipérazines	89
4)	Autres métabolites issus de la souche 497.....	91
E.	Induction de la production des métabolites d'actinobactéries	92
1)	Présentation	92
2)	Activité antibiotique	94
3)	Augmentation de l'activité de CoC1 et CoC2	96
Chapitre 2. Accéder à la diversité du microbiote bactérien cuticulaire.....		100

I.	Introduction	100
II.	Comparative analysis of DNA extraction methods to study the body surface microbiota of insects: a case study with ant cuticular bacteria.....	102
A.	Abstract	102
B.	Introduction.....	103
C.	Material and Methods	105
1)	Taxon sampling	105
2)	Overview of DNA extraction protocols	105
3)	Ant washing and DNA extractions.....	106
4)	DNA quantification	106
5)	PCR amplification and sequencing	106
6)	Sequence analysis	107
7)	Statistical analyses.....	108
D.	Results.....	110
1)	Comparison of DNA extraction protocols.....	110
(a)	Artifactual sequences and DNA contaminants	110
(b)	Effect on the number of reads, OTUs and diversity coverage.....	111
(c)	Taxonomic composition, and beta diversity	112
2)	Effect of pooling individuals and body size on cuticular diversity	114
E.	Discussion.....	116
1)	Effect of DNA extraction protocols on the cuticular bacterial diversity..	116
2)	Impact of number of individuals and individual body size	118
F.	Annexes : Figures et Tables du SI.....	120
	Chapitre 3. Jardin de fourmis : Entre microbiote et métabolome cuticulaire.....	121
I.	Introduction	121
II.	Partie expérimentale	124
A.	Prélèvements et préparation des échantillons	124
B.	Analyse par métabolomique	125

1) Acquisition des données	125
2) Contrôle de l'acquisition avec les QC	126
3) Préparation de la matrice de données pour les analyses multivariées	127
C. Séquençage haut débit du microbiote bactérien cuticulaire.....	128
1) Extraction des ADN, amplification 16S et séquençage haut débit.....	128
2) Prétraitements bioinformatiques.....	129
D. Analyses statistiques	130
1) Richesse en OTUs et en ions	130
2) OTUs partagés entre les deux espèces.....	131
3) Composition des communautés d'ions et d'OTUS	131
III. Résultats et discussions	133
A. Richesse en OTUs et en ions	133
B. OTUs partagés par les deux espèces.....	135
1) Résultats	135
1) Discussions	135
C. Composition des communautés d'OTUs	136
1) Analyses non supervisées	137
(a) Communautés bactériennes des sites.....	137
(b) Microbiote bactérien des cuticules	138
2) Analyses supervisées	138
(a) sPLS-DA et espèces.....	139
(b) sPLS-DA et site de prélèvement.....	141
D. Composition des communautés d'ions	143
1) Analyses non supervisées	143
2) Analyses supervisées :.....	144
(a) sPLS-DA et espèces.....	144
(b) sPLS-DA et site de prélèvement.....	148
E. Futures orientations.....	148

Conclusions générales et perspectives.....	150
Annexes	156
Bibliographie :	181

Avant-propos

En science, l'interdisciplinarité est le travail de recherche réalisé par un ensemble de chercheur-e-s et/ou d'équipes de recherches issus de disciplines scientifiques différentes. Le projet initial de cette thèse s'intitulait « Métabolites de fourmis de Guyane et de leurs micro-organismes associés : une approche d'écologie fonctionnelle comme source d'inspiration pour la découverte de composés anti-malariques et antimicrobiens ». Cependant, les objectifs scientifiques de ce travail ont été enrichis au cours de la thèse, notamment grâce à l'approche interdisciplinaire qui a toujours été mise en avant. Ainsi les trois chapitres de ce manuscrit s'articulent autour d'un thème commun « Microbiote bénéfique de la cuticule des fourmis » qui est abordé et étudié par différentes approches de biologie et de chimie. Le succès de cette stratégie repose sur des collaborations enrichissantes menées avec plusieurs chercheurs en biologie et en chimie de différentes spécialités incluant la microbiologie, la biologie moléculaire, l'écologie, la chimie des produits naturels et la chimie analytique. Le financement de la bourse doctorale, assurée par l'Institut National de Chimie du CNRS et par le Labex Centre d'Etude de la Biodiversité Amazonienne, et l'équipe encadrante, deux docteurs en chimie et deux docteurs en biologie, reflètent également l'interdisciplinarité de ce travail.

Introduction générale

La majorité des molécules antibiotiques actuellement utilisées en médecine sont issues directement ou indirectement de molécules produites par des actinobactéries (Baltz 2008). Cependant, la multi-résistance des souches bactériennes pathogènes chez l'homme à ces antibiotiques constitue un problème majeur de santé publique, et le besoin de découvrir de nouvelles molécules antibiotiques est plus que jamais de première importance (Davies & Davies 2010). Les actinobactéries se retrouvent dans les milieux marins ou terrestres au sein de communautés bactériennes complexes et extrêmement diversifiées avec lesquelles elles sont en interaction. Bien que le rôle écologique des antibiotiques produits par les actinobactéries reste indéterminé, il est maintenant admis qu'ils sont impliqués dans les interactions entre bactéries, et pourraient en particulier servir à la compétition pour les ressources.

Cependant, accéder à l'intégralité des métabolites produits *in vivo* représente un défi scientifique majeur. Premièrement, parce que l'analyse chimique de cultures de souches pures isolées de l'environnement ne permet pas de reproduire l'ensemble des cocktails antibactériens produit lorsque ces souches sont en interactions avec les autres membres de leur communauté. Néanmoins, il existe des stratégies de microbiologie classique qui pourraient augmenter les chances d'identifier des souches bactériennes et des structures chimiques originales. Par exemple, la coculture de différentes souches bactériennes permet, en induisant une compétition pour les ressources, d'induire la production de métabolites impliqués dans des mécanismes de défense ou de communication « inter-souche » ou « inter-spécifique » (Abdelmohsen *et al.* 2015) qui ne sont pas détectables en monoculture. Deuxièmement, la majorité des bactéries environnementales, y compris les actinobactéries, restent à ce jour non cultivables, ce qui restreint considérablement le champ de prospection des molécules bioactives. Néanmoins, les nouvelles technologies de chimie (métabolomique) et de biologie moléculaire (DNA metabarcoding et séquençage haut-débit) permettent maintenant d'étudier la diversité bactérienne et la diversité métabolique sur des échantillons environnementaux. Ces outils devraient donc améliorer notre compréhension des facteurs façonnant la composition des communautés bactériennes et le paysage chimique de la cuticule. Cette démarche interdisciplinaire pourrait donc à terme faciliter la prospection de produits antibiotiques nouveaux *in natura*.

Malgré ces avancées récentes, explorer le potentiel antibiotique des communautés bactériennes environnementales complexes reste un véritable challenge. Il est donc nécessaire de s'orienter vers des systèmes naturels simplifiés (Clardy *et al.* 2009). En cela, le microbiote, i.e. l'ensemble des micro-organismes (bactéries, levures, champignons, virus) associés à un hôte

(animal ou végétal) est un système plus accessible car il est moins diversifié et est un acteur fondamental dans l'écologie et l'évolution de son hôte. Les organismes supérieurs sont soumis tout au long de leur vie à la pression d'une multitude de prédateurs, de parasites et de pathogènes. Ils ont donc acquis de multiples traits mécaniques, chimiques, immunologiques et comportementaux liés à la défense. L'association symbiotique avec des micro-organismes protecteurs fait partie intégrante des processus évolutifs développés en réponse à cette pression extérieure. Des associations bénéfiques particulières sont de plus en plus décrites chez les plantes (Sheffer *et al.* 2015), les animaux (Kueneman *et al.* 2014), les insectes (Kaltenpoth & Engl 2014) et révèlent, notamment grâce à l'avènement des technologies de séquençage haut-débit, le rôle protecteur fondamental du microbiote pour son hôte. Cette protection se fait en partie par la production de molécules bioactives antimicrobiennes protégeant l'hôte contre des pathogènes en échange de ressources (Kaltenpoth 2009). L'interdépendance et la coévolution entre hôte et microbiote peut être si importante que des auteurs ont récemment proposé de ne plus considérer l'hôte et son microbiote comme des entités indépendantes mais plutôt comme une seule unité évolutive constituée de réseaux moléculaires biologiques de l'hôte et de ses micro-organismes associés appelée « holobionte » (Moran & Sloan 2015; Theis *et al.* 2016). Cependant, l'acquisition d'un microbiote bénéfique reste une question en suspens : il peut être transmis de manière verticale, ce qui peut mener à l'existence de souches – et potentiellement de molécules bioactives – co-évoluant avec leur hôte, mais il peut aussi être acquis dans l'environnement (Salem *et al.* 2015). Ceci appelle à une meilleure compréhension de la nature et de l'acquisition du microbiote ainsi que de déchiffrer son rôle dans le succès évolutif de son hôte.

Jusqu'à maintenant, la plupart des travaux s'intéressant à ce sujet sont principalement portés sur le microbiote intestinal de l'homme, ou plus généralement des vertébrés (Ley *et al.* 2008) et des insectes (Scott *et al.* 2008; Russell *et al.* 2009; Poulsen *et al.* 2011). Les microbiotes bactériens associés à la peau ou à la cuticule sont moins étudiés bien qu'ils jouent également un rôle important de défense contre les pathogènes en tant que première barrière entre l'hôte et son environnement (Grice & Segre 2011), notamment par le biais de la synthèse de substances antimicrobiennes. L'un des exemples terrestres les plus notables du mutualisme hôte-microbiote est celui des fourmis champignonnistes attines (Formicidae : Attini) et des actinobactéries cuticulaires avec lesquelles elles sont associées. Ces dernières participent activement à la défense de la colonie en synthétisant des composés fongicides contre les souches pathogènes (*Escovopsis*) du champignon cultivé par les attines pour leur consommation (Basidiomycètes) (Currie *et al.* 1999b). Le microbiote des fourmis représente un réservoir

inexploré et très prometteur pour découvrir de nouveaux antibiotiques. Ceci s'illustre par le nombre croissant de nouvelles structures bioactives isolées à partir d'actinobactéries issues du microbiote de certains insectes (Oh *et al.* 2009b; a, 2011; Gao *et al.* 2014). En outre, cette famille d'insecte constitue l'un des groupes les plus nombreux et écologiquement dominant d'animaux sociaux terrestres, avec des espèces retrouvées sur les six continents et une extrême diversité en milieu tropical (Basset *et al.* 2012). Leur succès évolutif, leur abondance et leur diversité s'expliquent en partie grâce aux interactions bénéfiques qu'elles entretiennent avec des micro-organismes (Moreau *et al.* 2006) qui leur ont permis de s'adapter à tous les environnements colonisés. Un exemple notable est celui de l'adaptation à la vie arboricole des fourmis du genre *Cephalotes*. En effet, ces fourmis herbivores ont un régime alimentaire pauvre en azote mais bénéficient d'un microbiote intestinal composé de bactéries assimilatrices d'azote (i.e. Rhizobiales) qui leur apporte les acides aminés nécessaires à leur bon développement physiologique (Russell *et al.* 2009). Les fourmis représentent un modèle idéal pour ce travail de bioprospection.

Ce travail de thèse a pour principaux objectifs (1) de caractériser la composition bactérienne et des produits naturels bioactifs potentiels de la cuticule des fourmis de Guyane, (2) de connaître quels sont les facteurs qui déterminent l'assemblage des communautés bactériennes et (3) d'identifier quels sont les taxons bactériens à l'origine de certains métabolites de la cuticule à potentiel bioactif. Il se situe à l'intersection de la microbiologie, la biologie moléculaire, l'écologie, la chimie des produits naturels et la chimie analytique et pourrait permettre de répondre à des questions fondamentales en écologie des communautés et en écologie chimique tout en contribuant à la recherche de produits d'intérêt pharmacologique. Ce travail de recherche est présenté dans ce manuscrit selon une organisation en trois chapitres.

Chapitre 1. Pouvoir antibiotique d'actinobactéries cuticulaires de fourmis

Le chapitre 1 est une étude en chimie des produits naturels qui s'intéresse aux activités antimicrobiennes d'actinobactéries isolées à partir des cuticules de diverses espèces de fourmis prélevées en Guyane. Ce chapitre s'inscrit d'une part dans une approche de prospection pharmacologique et d'autre part dans un contexte d'interaction biologique hôte-microbiote. En effet l'activité biologique des actinobactéries est évaluée contre des pathogènes humains ou de fourmis. Pour optimiser la chimiodiversité observée au laboratoire, ou pour simplement mieux représenter les interactions microbiennes des bactéries sur la cuticule d'une même fourmi, la mise en culture de plusieurs micro-organismes a été mise en place. Les techniques classiques d'isolement des métabolites bioactifs par fractionnement bioguidé et de détermination structurale ont été utilisées. De plus, à partir de mono- et de cocultures bactériennes et grâce

aux nouvelles technologies de traitement des données par spectrométrie de masse MS/MS permettant de générer des réseaux moléculaires pour mieux visualiser la diversité des métabolites, nous avons pu nous affranchir de l'isolement chronophage des métabolites.

Chapitre 2. Accéder à la diversité du microbiote bactérien cuticulaire

Le chapitre 2 décrit l'étude métabarcoding du microbiote bactérien cuticulaire de deux espèces de fourmis : *Atta cephalotes* et *Pseudomyrmex penetrator*. La composition des communautés bactériennes analysée par séquençage n'a pas fait l'objet de nombreuses études jusqu'à présent et avant de comprendre les rôles biologiques de ces communautés, il est nécessaire dans un premier temps de mettre en place un protocole efficace pour accéder à la diversité bactérienne d'échantillons de petite taille. Ce travail représente un défi technique important en écologie des communautés de bactéries associées aux insectes. Le microbiote cuticulaire des échantillons a été extrait par quatre méthodes différentes dans le but d'identifier le meilleur protocole d'extraction d'ADN (Acide DésoxyriboNucléique) bactérien à partir de la cuticule d'un seul individu. Les effets du nombre d'individus poolés dans un même échantillon ou de la taille des individus collectés ont été également adressés dans ce chapitre. Le contenu de ce chapitre a fait l'objet d'un article soumis au journal *Molecular Ecology Resources* en décembre 2016 et est actuellement en cours de révision. Cet article correspond au chapitre 2 de ce manuscrit de thèse et il a été par conséquent intégré en anglais.

Chapitre 3. Jardin de fourmis, entre microbiote et métabolome cuticulaire

Le chapitre 3 est une étude inédite qui combine l'interprétation de données de métabolomique avec des données de métabarcoding à partir de la cuticule de deux espèces de fourmis vivants ensemble (i.e. parabiose) dans des jardins de fourmis. Ces deux espèces de fourmis sont *Camponotus femoratus* (*Ca. femoratus*) et *Crematogaster levior* (*Cr. levior*). L'originalité de cette étude est la mise en lumière de nouveaux éléments permettant d'une part d'appréhender les mécanismes de transmission des communautés bactériennes cuticulaires des fourmis et d'autre part de tenter d'accéder aux métabolites du microbiote bactérien cuticulaire. Cette étude du microbiote en parallèle de celle du métabolome cuticulaire de deux espèces de fourmis a été réalisée sur plusieurs nids collectés à une échelle régionale.

Etat de l'art général des modèles d'étude

I. Ecologie des communautés microbiennes

A. Généralités

1) *Communautés microbiennes, « espèce » et fonctions*

Les communautés microbiennes sont des assemblages de plusieurs « espèces » de micro-organismes (bactérie, champignon, levure, virus) vivant dans un environnement contigu et interagissant les uns avec les autres. L'écologie microbienne est une discipline qui cherche à (i) comprendre les facteurs qui régissent les assemblages microbiens dans l'espace et le temps (ii) appréhender l'implication fonctionnelle de ces communautés dans leur environnement biotique et abiotique afin de (iii) mieux prédire l'évolution de ce compartiment biologique dans un environnement changeant. Notons que dans le cadre de cette thèse, nous nous intéressons uniquement aux communautés bactériennes.

Le concept d'espèce est nécessaire à la compréhension et à la connaissance de la composition et de la structure des communautés bactériennes. Chez les procaryotes, la notion d'espèce est un concept difficilement définissable et les microbiologistes ne sont pas encore parvenus à un consensus pour définir la plus petite unité fondamentale de la diversité bactérienne (Ward *et al.* 2008). D'un point de vue moléculaire, il est admis que deux souches présentant plus de 3% de différence dans leurs séquences d'ARN (Acide RiboNucléique) ribosomal appartiennent systématiquement à des espèces différentes (Tindall *et al.* 2010). Les limites de cet indice et le besoin d'intégrer la dimension écologique et évolutive ont permis de développer le concept d'écotype. L'écotype est le concept d'espèce écologique, c'est à dire une espèce composée d'individus considérés identiques dans toutes leurs propriétés écologiques (Cohan 2002). Il s'agit d'une population de bactéries occupant la même niche écologique discrète, dont la variation génétique est purgée de façon récurrente par la sélection naturelle. De plus, l'importance des recombinaisons homologues et des transferts de gènes horizontaux varie selon les bactéries étudiées, ce qui constitue une autre explication de la difficulté à obtenir un consensus sur l'espèce bactérienne (Prosser *et al.* 2007).

Dans le cadre de ce manuscrit, nous utilisons le terme de souche bactérienne pour les études de microbiologie où des actinobactéries sont isolées. Dans le cadre de l'étude métabarcoding présentée, nous utilisons le terme OTU pour Operational Taxonomic Unit. Un OTU est la séquence la plus abondante d'un groupe de séquences semblables à 97% et est assimilé à un taxon via une base de données lorsque le degré d'assignation est supérieur à 80%.

Les fonctions biologiques sont constituées de l'ensemble des activités physiologiques, biochimiques et physiques présentes à l'échelle de l'organisme ou de la communauté. Ces fonctions influencent de manière directe ou indirecte non seulement le porteur de la fonction mais également son environnement (Madigan & Martinko 2007). Au niveau de l'organisme, les fonctions sont le résultat combiné de l'expression du génome, des conditions environnementales et relèvent principalement du métabolisme. Au niveau de la communauté, la combinaison des fonctions en donne de nouvelles qui participent au fonctionnement de l'écosystème. L'étude des fonctions de la communauté bactérienne est possible via l'étude des génomes de la communauté (métagénome), des gènes traduits dans la communauté (métatranscriptome) ou des protéines synthétisées par la communauté (métaprotéome) : les outils permettant d'analyser ces ensembles sont détaillés dans l'état de l'art des méthodes de ce manuscrit.

2) *Distribution et règles d'assemblage*

« *Everything is everywhere, but, the environment selects* » (Bass-Becking 1934).

L'écologie des communautés s'intéresse aux règles d'assemblage des communautés, c'est à dire aux processus qui déterminent la composition et la structure d'une communauté de plantes, d'arthropodes ou de microorganismes. La citation ci-dessus est célèbre en microbiologie ; elle constitue l'une des premières hypothèses d'écologie des communautés chez les micro-organismes. La notion portée par « *everything is everywhere* » fait référence à la capacité apparemment illimitée de dispersion des micro-organismes. Cette dispersion se fait de manière passive (courants, vents, animaux vecteurs) et est accrue par un taux de reproduction élevé (Finlay & Clarke 1999; Martiny *et al.* 2006). La notion portée par « *the environment selects* » s'apparente à la théorie de la niche et souligne le fait que certaines espèces microbiennes sont adaptées à un milieu donné. C'est la notion de filtre environnemental. L'hypothèse de Baas-Becking revient donc à dire que la distribution spatiale (ou temporelle) des taxons microbiens est uniquement fonction des conditions abiotiques. Cependant, l'hypothèse de Baas-Becking ne prend pas en compte les interactions entre les micro-organismes, par exemple la compétition (Ghoul & Mitri 2016). D'autre part, la distribution des bactéries peut être aussi contrôlée par des processus aléatoires, lorsque la dispersion est limitée. Par exemple, la composition de deux communautés microbiennes peut se différencier à court terme via des processus démographiques stochastiques (e.g. naissance, mort, (Curtis & Sloan 2004; Caruso *et al.* 2011)). Dans ce cas, les taxons de la communauté microbienne peuvent être vus comme étant tirés au sort au sein du pool d'espèces initial. A long terme, chacun des taxons constituant la communauté peut aussi évoluer de façon neutre (mutation aléatoire, (Hanson *et*

al. 2012)). L'importance relative du filtre environnemental et des processus aléatoires semble donc déterminée par le taux de dispersion des microorganismes. Bien que certaines études semblent suggérer que la dispersion microbienne n'est pas illimitée, contrairement à ce que sous-entend l'hypothèse de Baas-Becking, il reste actuellement difficile de le valider – ou tout du moins de le généraliser – car les outils permettant d'identifier les taxons bactériens dans l'environnement ne sont pas encore assez sensibles pour discriminer des entités bactériennes phylogénétiquement proches (e.g. équivalents de souches, ou d'espèces ; (Fuhrman 2009)).

B. Les interactions hôtes-micro-organismes : Le microbiote

1) Définitions

Le microbiote est l'ensemble « d'espèces » microbiennes (englobant bactérie, levure, champignon, virus) vivant dans un environnement spécifique (appelé microbiome) chez un hôte (animal ou végétal). Le microbiome est une extension du microbiote ; il correspond à l'ensemble des gènes (ou métagénome) du microbiote. On qualifie de symbiose toute association entre différents organismes, ici microbiote et hôte. Bien que le terme soit souvent employé pour définir une association positive, ce terme ne présuppose aucune modalité d'interaction.

La symbiose a été pour la première fois décrite par Antoine de Bary en 1879 pour décrire deux ou plusieurs espèces différentes partageant une association intime et durable (de Bary 1879; Oulhen *et al.* 2016). On distingue différentes catégories de symbiose selon la modalité de l'interaction entre le micro-organisme symbiotique et son hôte. Ainsi, si la symbiose est bénéfique pour les deux parties on emploie le terme de mutualisme, si elle bénéficie à une partie et n'affecte pas l'autre on emploie le terme de commensalisme et si elle est au détriment d'une des parties et au bénéfice de l'autre on emploie le terme de parasitisme. On différencie également les micro-organismes symbiotiques facultatifs, qui ne sont pas nécessaire à la survie de l'hôte, et les micro-organismes symbiotiques obligatoires qui sont indispensables à sa survie (Brownlie & Johnson 2009). Dans de nombreux cas, la symbiose est caractérisée par une association spécialisée entre un hôte et un micro-organisme spécifique à l'échelle de l'espèce, qui lui fournit une ou plusieurs fonctions bénéfiques. Parmi elles, on peut citer l'acquisition de voies métaboliques lui permettant d'accéder à un nouvel apport nutritif, ou bien encore à un rôle de défense contre les pathogènes (Moran 2006), qui nous intéresse tout particulièrement dans le cadre de cette thèse. Au sein de ces associations, les avantages apportés par les micro-organismes symbiotiques peuvent être contextuellement dépendants et influencés par d'autres taxons microbiens associés à l'hôte ou présents dans l'environnement. En outre, les symbioses dans la nature sont plus complexes. Elles peuvent par exemple impliquer plusieurs micro-organismes

qui vont agir de manière coordonnée, concurrentielle ou indépendante sur la « *fitness* » de l'hôte (Hussa & Goodrich-Blair 2013). On parle dans ce cas de symbiose multipartite.

La théorie hologénomique de l'évolution des systèmes symbiotiques hôte-microbiote postule que le devenir évolutif du microbiote est lié à celui de l'hôte (Rosenberg & Zilber-Rosenberg 2016; Theis *et al.* 2016). En d'autres termes, les forces évolutives agissent sur l'holobionte (ou son génome, l'hologénome), qui est une unité biologique composée de l'hôte et de son microbiote (ou de la somme des génomes de l'hôte et de son microbiote). Cette théorie est au cœur des débats et s'oppose à une autre théorie qui consiste à penser que l'hôte et son microbiote sont des communautés d'organismes qui englobent un large panel d'interactions (du parasitisme au mutualisme), des modes de transmission rapportés (horizontale à verticale) ainsi que différents niveaux de fidélité parmi les partenaires (Douglas & Werren 2016). La manière dont le microbiote évolue sous la sélection naturelle pour être bénéfique à l'hôte pourrait être plus complexe que celle proposée par la théorie de l'hologénome (Moran & Sloan 2015). Ces deux visions sont différentes mais pas exclusives, il se pourrait qu'un juste milieu soit plus proche des réalités.

2) *Acquisition et transmission*

L'étude de l'acquisition et de la transmission des microorganismes symbiotiques s'apparente à l'écologie des communautés (cf section précédente). Les micro-organismes symbiotiques bénéfiques sont transmis entre générations par voie verticale, directement des parents à la progéniture. Ils peuvent être acquis par voie horizontale, indirectement entre les individus d'une même espèce ou via l'environnement ou encore par une combinaison des deux voies (Dieter 2013). Il est important de distinguer les micro-organismes symbiotiques intracellulaires qui sont le plus souvent impliqués dans des relations symbiotiques obligatoires et transmis par voie maternelle (cellules ovariennes), des micro-organismes symbiotiques extracellulaires qui eux ont de multiples voies de transmission chez les insectes par exemple (Salem *et al.* 2015). Le point commun des micro-organismes symbiotiques extracellulaires est leur capacité à survivre à l'extérieur de leur hôte contrairement aux organismes intracellulaires pour qui cette faculté s'est perdue au cours la coévolution hôte-micro-organisme (Moran *et al.* 2008). Récemment six voies de transmission différentes des micro-organismes symbiotiques extracellulaires ont été proposées en s'appuyant sur de nombreux exemples chez les insectes, ce qui est plus précis que la distinction transmission verticale et horizontale. Ainsi, on distingue l'acquisition dans l'environnement, la coprophagie, l'inoculation sur les œufs, l'acquisition sociale, la transmission par capsule et la contamination via des sécrétions. Chez les insectes de l'ordre des Hyménoptères, seule la coprophagie, l'inoculation sur les œufs et la transmission

sociale sont les modes de transmission de micro-organismes symbiotiques extracellulaires rapportés dans la littérature (Salem *et al.* 2015).

3) *Exemples de systèmes hôte-microbiote*

Grâce au développement des technologies de séquençage haut-débit, de nouvelles méthodes en écologie moléculaire permettent maintenant d'explorer le monde microbien à un niveau de résolution jamais atteint. Ainsi il est possible d'étudier à moindre coût les interactions complexes entre l'hôte et son microbiote (Ley *et al.* 2008; Knight *et al.* 2012; Shinzato *et al.* 2014). Dans ce contexte plusieurs projets internationaux de grande envergure sont en cours de réalisation. Le Human Microbiome Project (<http://hmpdacc.org/overview/about.php>) est un projet débuté en 2008 dont l'objectif est d'identifier et de caractériser les micro-organismes trouvés chez des humains sains et malades (Turnbaugh *et al.* 2007; Peterson *et al.* 2009). Il a permis – entre autres – le développement de nouvelles thérapies pro- ou antimicrobiennes (Vincent & Manges 2015). Par exemple, des gènes bactériens codant pour la biosynthèse de produits naturels ayant de remarquables activités biologiques contre des micro-organismes multi résistants ont été identifiés chez deux bactéries commensales des humains (Zipperer *et al.* 2016; Chu *et al.* 2016). Ces études se généralisent maintenant sur des systèmes non modèles (Ley *et al.* 2008; Russell *et al.* 2009).

La surface du corps des vertébrés et invertébrés abrite des communautés microbiennes diverses qui ont été beaucoup moins étudiées par rapport au microbiote intestinal. Les communautés microbiennes de surface sont impliquées dans les mécanismes de défense de l'hôte par le blocage physique des agents pathogènes (production de biofilms microbiens), par la compétition nutritive avec les agents pathogènes ou par la biosynthèse de métabolites antimicrobiens (Brownlie & Johnson 2009). Les mécanismes d'acquisition du microbiote cutané/cuticulaire restent encore largement méconnus, mais il semblerait qu'il diffère entre espèces d'hôtes, comme le suggère une étude du microbiote cutané d'amphibiens (Kueneman *et al.* 2014). Cependant, l'effet espèce peut refléter des différences d'habitats, de distribution géographique, ou d'histoires co-évolutives, et nécessite d'être étudié plus en avant.

II. Les actinobactéries

Dans le cadre de ce travail de thèse, nous porterons une attention particulière aux actinobactéries, car elles sont à l'origine de la majorité des molécules antibiotiques actuellement utilisées en médecine (Baltz 2008). Les substances naturelles produites par les actinobactéries constituent toujours une source importante de molécules antibiotiques, cependant certains écueils tel l'isolement de molécules déjà connues nécessite la mise en place de nouvelle stratégie pour mieux explorer la diversité des actinobactéries et de leur métabolome (Baltz 2008).

A. Présentation

Les actinobactéries sont des bactéries filamenteuses, dont le membre le plus connu, *Streptomyces sp.*, présente une grande diversité en termes de morphologie, de physiologie et de capacités métaboliques (Barka *et al.* 2016). Les actinobactéries (anciennement appelées actinomycètes), littéralement « champignon rayonnant », sont un groupe d'eubactérie Gram-positives, avec un fort taux de guanine+cytosine (G+C) dans leur ADN. Elles ont un cycle biologique complexe qui comprend la formation de cellules filamenteuses appelées hyphes qui se différencient pour donner des spores asexuées (Figure 1) (Barka *et al.* 2016). Sur une gélose solide, les actinobactéries développent un réseau ramifié d'hyphes poussant à la fois à la surface et à l'intérieur pour former le mycélium végétatif (Prescott *et al.* 2010) (Figure 1).

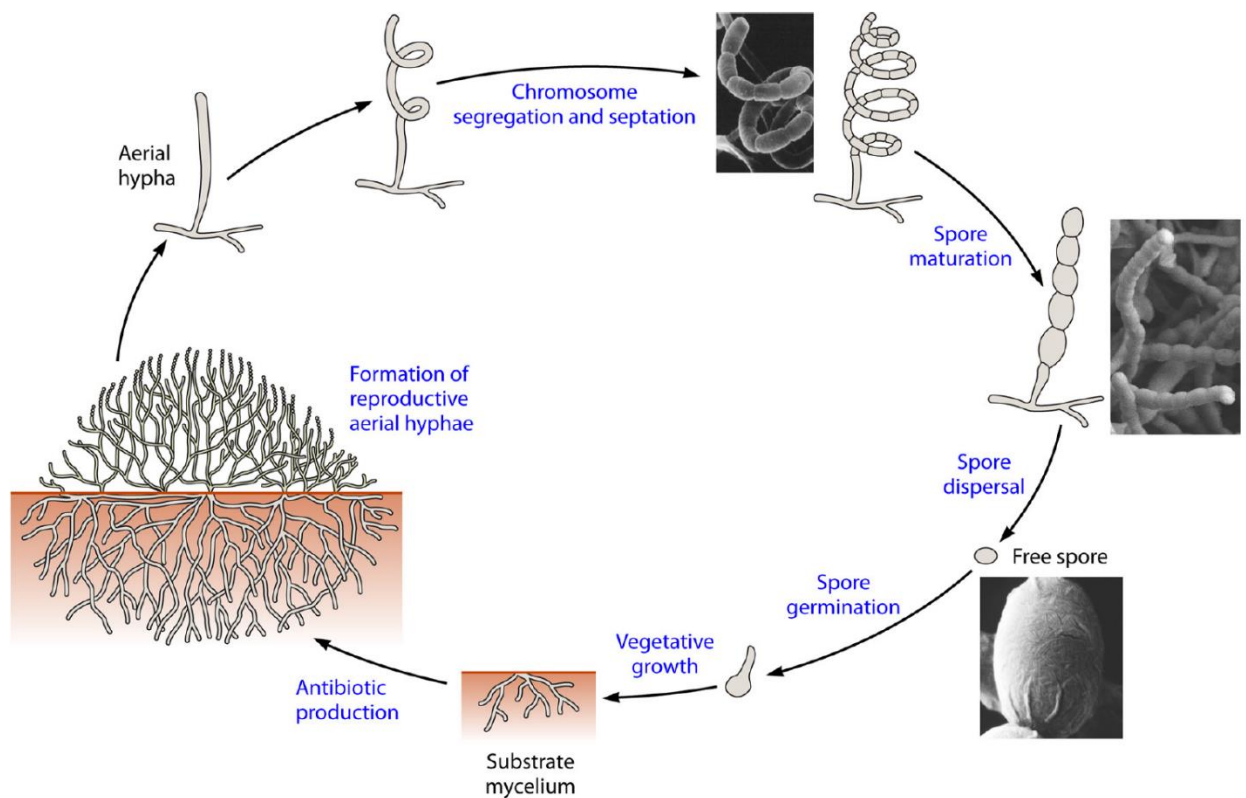


Figure 1 : Schéma du cycle de vie des actinobactéries incluant un schéma d'une coupe transversale d'actinobactérie avec des hyphes vivants (Barka *et al.* 2016).

En utilisant des critères moléculaires et chimiques, on discrimine actuellement 12 sous-ordres, 39 familles et 130 genres (Figure 2) (Barka *et al.* 2016). On retrouve les actinobactéries dans les écosystèmes aquatiques et terrestres, bien qu'elles soient plus abondantes dans les sols que dans les autres milieux, en particulier dans les sols alcalins et dans les sols riches en matière organique (Prescott *et al.* 2010). Les actinobactéries peuvent être trouvées à la fois sur la surface du sol et à des profondeurs de plus de deux mètres (Goodfellow & Williams 1983). Elles jouent un rôle clé dans l'écologie des sols en permettant la minéralisation de la matière organique. En particulier, elles participent à l'hydrolyse d'une large gamme de polysaccharides (cellulose, chitine, xylane et agar) et d'autres macromolécules naturelles (Chater *et al.* 2010). De nombreuses actinobactéries sont de nos jours isolées des écosystèmes marins et ces souches présentent une physiologie unique leur permettant de survivre à des conditions extrêmes (pression, salinité et basses températures). Nombre d'entre elles produisent des molécules bioactives originales jamais isolées chez leurs homologues terrestres (Abdelmohsen *et al.* 2014).

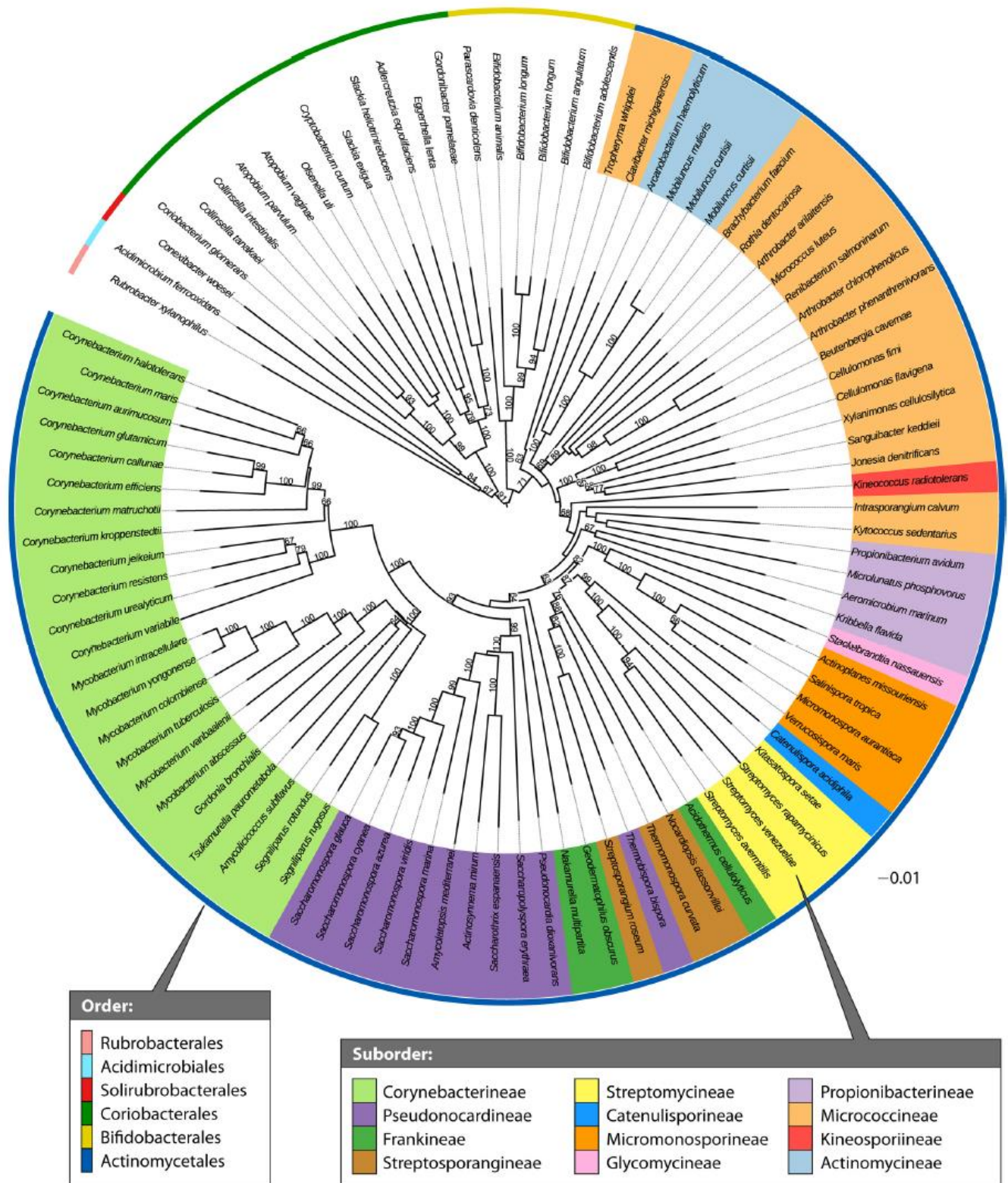


Figure 2 : Phylogénie moléculaire de 97 séquences du génome du phylum Actinobacteria (Barka et al. 2016).

B. Intérêts pharmaceutiques

Les actinobactéries ont la capacité de produire une grande diversité de produits naturels possédant une large gamme d'activités biologiques. Par exemple de nombreux composés antibiotiques, antifongiques, antiparasitaires, antiviraux, anti-cellules tumorales, herbicides, pesticides ont été isolés depuis les années 1950 (Figure 3)(Barka *et al.* 2016). Parmi ces activités biologiques, les molécules antibiotiques produites par les actinobactéries sont particulièrement importantes en médecine et incluent les aminosides, les anthracyclines, le chloramphénicol, les macrolides et les tétracyclines (Figure 3) (Barka *et al.* 2016). Au sein des actinobactéries, plus de 10 000 molécules bioactives sont connues, dont 7600 dérivent des espèces du genre *Streptomyces* et 2500 d'espèces plus rares d'actinobactéries (Bérdy 2005). Ces bactéries filamenteuses sont responsables de la production d'environ 70% des antibiotiques utilisés en pharmacie, ce qui fait de ce groupe la première source d'antibiotiques (Bérdy 2005). La streptomycine isolée de *Streptomyces griseus* (Figure 4) est le premier antibiotique de la classe des aminoglycosides efficace contre la tuberculose et découvert en 1943 par Albert Schatz, étudiant à l'époque dans un laboratoire public dans le cadre d'une collaboration avec l'entreprise pharmaceutique Merck (Kingston 2004). Après l'âge d'or de la découverte des antibiotiques dans les années 1950-1960, les années 1980 ont connues un obstacle de taille avec la redécouverte de produits naturels d'actinobactéries ce qui a entraîné dans les années 2000 une diminution des efforts pour isoler de nouvelles molécules antibiotiques. Ce phénomène, additionné à celui de l'augmentation des résistances des micro-organismes pathogènes aux antibiotiques commercialisés (Blair *et al.* 2015), démontre le besoin nécessaire de découvrir de nouvelles molécules antibiotiques originales en utilisant de nouvelles stratégies en biologie et en chimie.

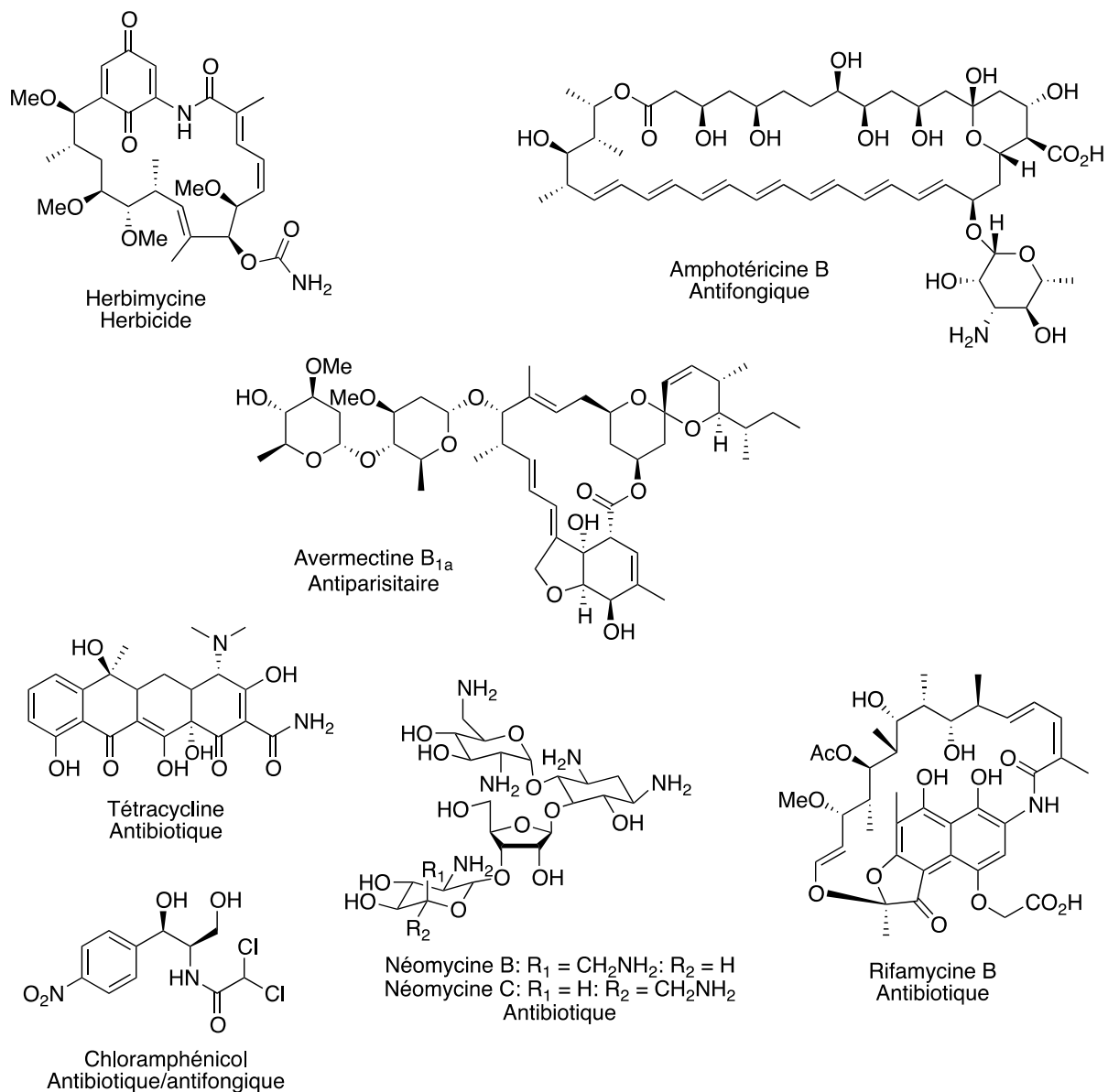


Figure 3 : Exemples de molécules actives issues d'actinobactéries (Barka *et al.* 2016).

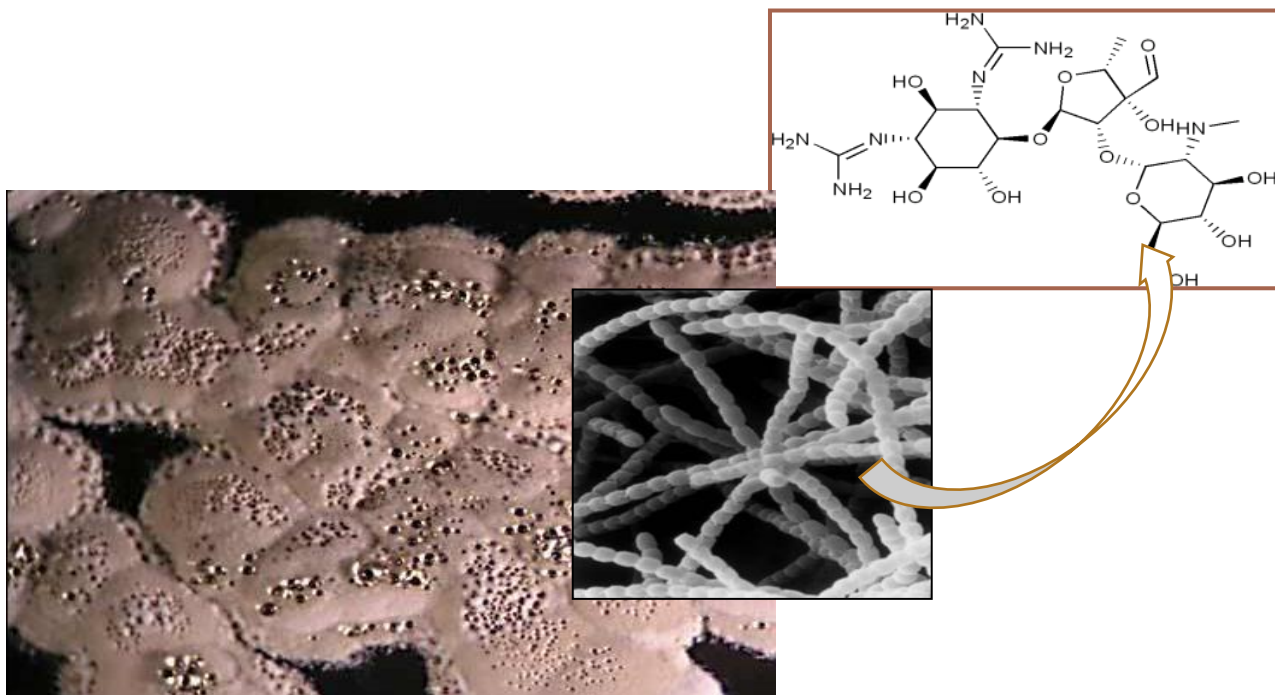


Figure 4 : *Streptomyces griseus* et la molécule de streptomycine.

A l'heure actuelle, les produits naturels d'actinobactéries bénéficient d'un regain d'attention grâce à l'acquisition massive de données de séquences d'actinobactéries et le développement de nouvelles méthodes permettent maintenant la manipulation génétique ou physiologique (Abdelmohsen *et al.* 2015) de souches connues, l'isolement de souches à partir de sources nouvelles (Challinor & Bode 2015), des approches de cocultures (Bertrand *et al.* 2014), des études de génome mining (Ju *et al.* 2015), des corrélations de cluster de gènes avec de la métabolomique (Goering *et al.* 2016) et la culture in situ des souches « non-cultivables » selon la technique du iChip (Ling *et al.* 2015). Ces avancées suggèrent que le génome des actinobactéries contient des gènes codant pour la production— ou la résistance à — de multiples produits naturels n'ayant jamais été observés au laboratoire avec des techniques de culture simple (Baltz 2008; Harvey *et al.* 2015).

C. Mutualisme entre actinobactéries cuticulaires et leurs hôtes

Les actinobactéries sont souvent associées à des hôtes, notamment dans leur microbiote cutané/cuticulaire, avec lesquels elles entretiennent une relation mutualiste. En particulier, elles sont impliquées dans la production d'antibiotiques qui protègent l'hôte d'attaque de pathogènes (Kaltenpoth 2009). On retrouve des interactions positives microbiote cutané/cuticulaire-hôte notamment chez les hommes (Oh *et al.* 2014), les amphibiens (Kueneman *et al.* 2014), et les insectes (Kaltenpoth & Engl 2014). L'un des exemples les plus notables du mutualisme hôte-microbiote cutané/cuticulaire est celui des fourmis champignonnistes de la tribu des attines, détaillé dans la partie suivante consacrée au microbiote cuticulaire des fourmis (Currie *et al.*

1999b). Il est intéressant de noter que la mise en culture des actinobactéries cuticulaires d'insectes a permis d'isoler de nouvelles structures bioactives contre des pathogènes humains (Figure 5) (Oh *et al.* 2009b; a, 2011; Gao *et al.* 2014).

Chez les insectes, et plus particulièrement chez les Hyménoptères, d'autres exemples d'associations positives avec des actinobactéries ont été rapportées. Les guêpes solitaires construisent leur nid dans le sol pour élever leurs larves qu'elles nourrissent avec des insectes morts. Ce comportement constitue un risque d'infection élevé en exposant les larves et les guêpes à de nombreux pathogènes extérieurs au nid. Pour se prémunir de telles infections elles exécutent régulièrement des mouvements de la tête qui mettent en contact leurs antennes avec la surface du nid et les larves. En effet leurs antennes présentent des accrétions macroscopiques d'actinobactéries productrices d'antibiotiques qu'elles déposent par contact (Kroiss *et al.* 2010). Plus de 200 souches de *Streptomyces* ont été isolées à partir des lavages de la cuticule des guêpes solitaires et les analyses chimiques par HPLC/MS des cultures bactériennes ont permis de détecter des molécules antibiotiques : Stretazoline, Streptazone B, Bafilomycine A1, Bafilomycine B1, Daunomycine, Antimycine (A1 à A4) ainsi qu'une molécule antifongique décrite pour la première fois, Sceliphrolactam, active contre une souche *Candida albicans* résistante à l'amphotéricine B (Figure 5). Par analogie avec l'association entre actinobactéries et fourmis de la tribu des attines, l'hypothèse du rôle défensif des actinobactéries chez les guêpes solitaires est supportée par la diversité moléculaires des métabolites bioactifs biosynthétisés par des communautés bactériennes du genre *Streptomyces* (Oh *et al.* 2011; Poulsen *et al.* 2011). La diversité de métabolites impliquent un rôle de défense car de manière parallèle à la combinaison de plusieurs molécules en médecine humaine, le cocktail de molécules antibactériennes produites par la communauté de *Streptomyces* permet de diminuer l'apparition de résistances, par rapport à une situation propice à l'apparition de résistances où il y aurait un seul antibiotique (Kroiss *et al.* 2010).

D'autres exemples de molécules issues d'actinobactéries de la cuticule d'insectes ont été décrits. Le Microtermolide A et la Mycangimycine ont été isolés à partir de *Streptomyces* de la cuticule de termites et de coléoptères (*Dendroctonus frontalis*) respectivement, et présentent des activités biologiques intéressantes de type antifongique et antimalarique pour la Mycangimycine (Figure 5) (Oh *et al.* 2009b; Carr *et al.* 2012).

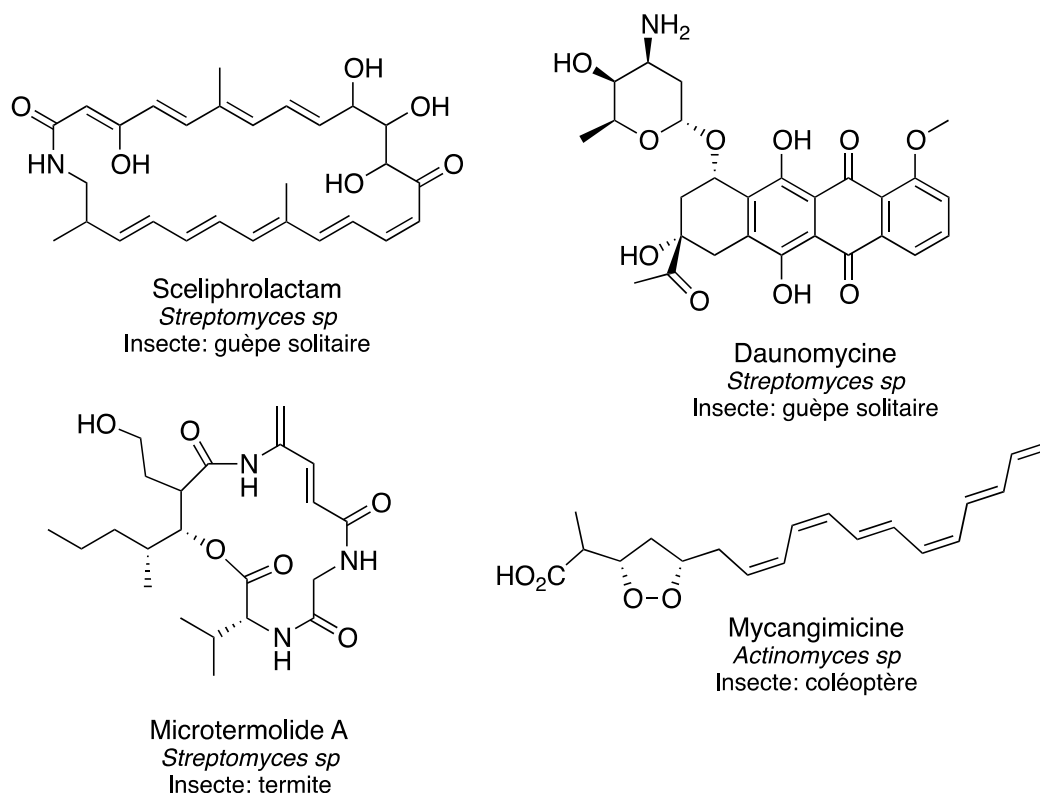


Figure 5 : Structure chimique de molécules isolées à partir de la culture d'actinobactéries collectées sur la cuticule d'insectes.

III. Le microbiote bactérien de la cuticule des fourmis

A. Contexte

Les interactions entre les fourmis et les micro-organismes présents dans leur habitat sont complexes. La promiscuité entre les individus, l'environnement riche en micro-organismes et l'apport d'éléments étrangers à la fourmilière en son sein (proies, feuilles...) constituent des risques infectieux pour la colonie. Cependant leur diversité d'espèces et leur abondance suggère que les stratégies de défense contre les pathogènes des fourmis sont extrêmement performantes. Les principaux éléments de ce succès s'expliquent par leur système immunitaire performant (travaux de Jules Hoffman (Lemaitre *et al.* 1996)), les mécanismes d'hygiène (zone déchets, lavages/léchages « grooming » entre individus) et les associations avec des micro-organismes sur leur cuticule à but défensif. Dans ce sens, la cuticule des fourmis est à la fois une barrière physique contre l'intrusion de micro-organismes dans l'hémolymphe et un substrat pouvant être colonisé par de nombreux micro-organismes bénéfiques (Douglas 2015). L'interaction positive microbiote cuticulaire-fourmis a même conduit au développement, par processus de coévolution, à des structures morphologiques particulières chez l'hôte. On peut citer l'exemple des cryptes retrouvées sur la cuticule des attines hébergeant des actinobactéries (Currie *et al.*

2006). Certaines fourmis présentent aussi une glande métapleurale, qui assure le maintien de l'hygiène chez l'hôte (Yek & Mueller 2011). Elle correspond à une invagination de la cuticule où sont stockées des sécrétions acides aux propriétés antibiotiques qui se répandent passivement par un orifice continuellement ouvert.

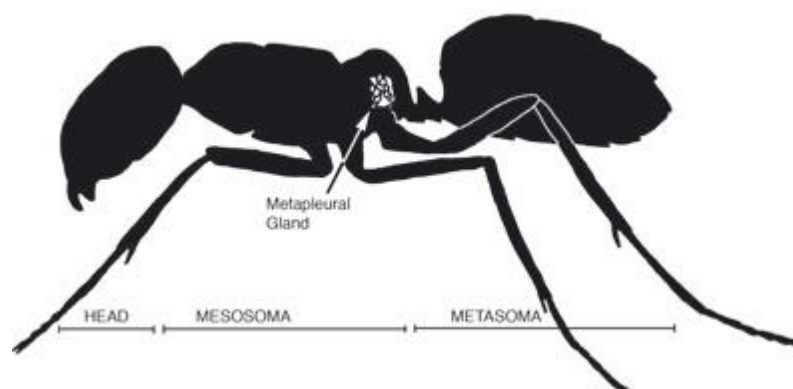


Figure 6 : Position de la glande métapleurale (Yek & Mueller 2011).

B. Les attines

De très nombreuses études se sont intéressées aux communautés d'actinobactéries associées aux attines. Les fourmis attines d'Amazonie cultivent depuis 45 millions d'années des champignons de la famille des Basidiomycètes comme principale source d'alimentation pour l'élevage des larves. Ce modèle représente un modèle de coévolution remarquable entre des insectes et des micro-organismes (Chapela *et al.* 1994). La souche fongique est transmise de façon verticale aux nouvelles générations et cette transmission est assurée par les reines vierges qui transportent du mycélium dans leurs poches infra-buccales. Cependant, cette symbiose est menacée par des champignons pathogènes (*Escovopsis spp.*) qui attaquent le cultivar des fourmis (Currie *et al.* 1999a). Plusieurs stratégies de défense contre ces pathogènes ont été mises en évidence : le champignon cultivé est capable de produire des molécules de défense (Wang *et al.* 1999) et les fourmis « désherbent » leur cultivar en enlevant les spores de champignon étrangers à leurs cultures avec leurs poches infra buccales pour les déposer en dehors du nid (Little *et al.* 2006). Enfin, les attines utilisent les métabolites antimicrobiens produits par leur glande métapleurale (Bot *et al.* 2002) et par les actinobactéries présentes sur leur cuticule.

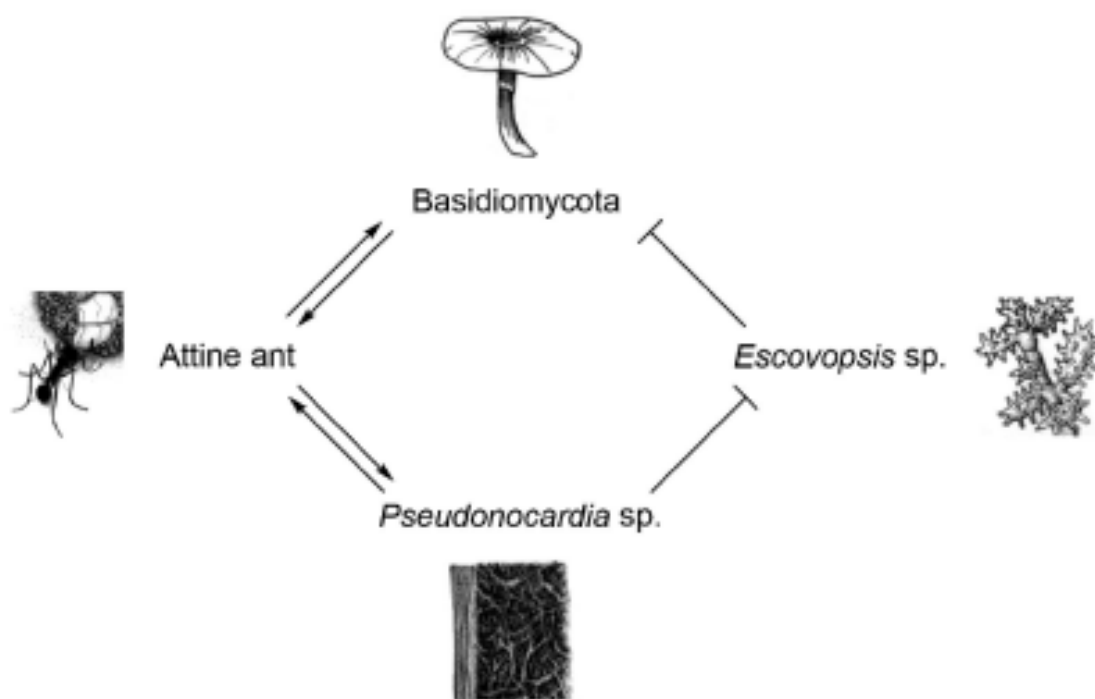


Figure 7 : Mutualisme des attines (Ramadhar *et al.* 2014).

Ce dernier aspect a fait l'objet de beaucoup d'attention suite à la première identification d'une molécule bioactive d'origine bactérienne. La dentigérumycine est un peptide non-ribosomique cyclique présentant des activités antifongiques et antibiotiques. Elle est produite par les actinobactéries du genre *Pseudonocardia* isolées sur la cuticule des attines de l'espèce *Acromyrmex octospinosus* (Currie *et al.* 1999b; Oh *et al.* 2009a). Depuis, d'autres souches de *Pseudonocardia* ont été isolées et de nouvelles structures chimiques bioactives ont été caractérisées (Sen *et al.* 2009). Des macrolides de type candicidine antifongiques ont également été détectés chez des *Streptomyces* (Haeder *et al.* 2009) (Figure 8).

Ces premières études font l'hypothèse que les bactéries du genre *Pseudonocardia* ont co-évolué avec leurs hôtes et sont transmises verticalement par les reines avec la souche fongique symbiotique. Cependant, d'autres études montrent que les attines sont associées avec des actinobactéries du genre *Streptomyces* et *Amycolatopsis* qui seraient acquises horizontalement à partir du sol ou pendant la phase de dispersion des mâles (Mueller *et al.* 2008). De plus, bien que certaines molécules antifongiques, comme la dentigérumycine, suggèrent une transmission verticale des souches du genre *Pseudonocardia* possiblement associée à un processus de coévolution, d'autres molécules antifongiques isolées et bien connues parmi les bactéries environnementales suggèrent une acquisition des souches dans l'environnement.

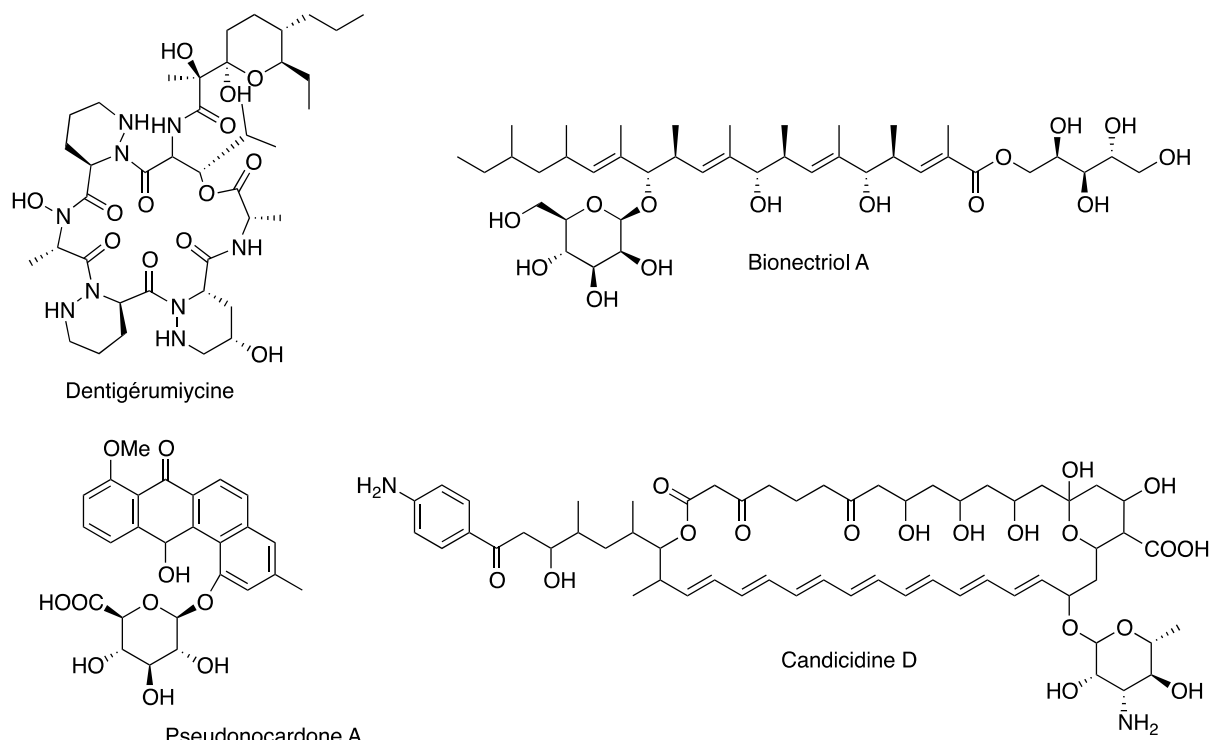


Figure 8 : Molécules actives isolées à partir d'actinobactéries de la surface de fourmis attines.

C. Autres fourmis

Trois symbioses mutualistes impliquant des fourmis et leur plante hôte ont été étudiées pour leurs actinobactéries cuticulaires : *Pseudomyrmex penetrator* et *Tachigali sp* (Fabaceae) en Guyane, *Petalomyrmex phylax* et *Leonardoxa africana* (Leonardoxa) et *Crematogaster margaritae* et *Keetia hispida* (Rubiaceae) au Cameroun. Ces fourmis arboricoles cultivent des souches fongiques pour nourrir leurs larves ce qui laissait présager d'un rôle défensif protecteur des actinobactéries pour les fourmis et leur cultivar (Hanshew *et al.* 2015). Les actinobactéries cuticulaires isolées à partir de ces espèces ont été mises en culture et montrent bien une activité d'inhibition de la croissance de souches pathogènes.

Deux autres symbioses mutualistes impliquant des fourmis myrmécophytes et leur plante hôte ont été étudiées. Dans ce travail, la diversité des communautés bactériennes cuticulaires a été évaluée par séquençage haut-débit pour la première fois (Seipke *et al.* 2013). La première association est composée par les fourmis *Allomerus spp.* vivant dans *Hirtella physophora* (Chrysobalanaceae) qui cultivent une souche fongique pour rigidifier les galeries déjà existantes de la plante dont elles se servent pour piéger leurs proies dans un but nutritif (Dejean *et al.* 2005). Dans la deuxième association, les fourmis *Tetraponera spp* vivent dans *Acacia drepanolobium* (Fabaceae) et cultivent les champignons comme source nutritive. Le microbiote bactérien cuticulaire de ces deux espèces de fourmis, qui sont phylogénétiquement

proches, comprend plus de 75% de Protéobactéries (principalement Gammaprotéobactéries) et la composition des phylums bactériens majoritaires est comparable bien que ces deux espèces vivent sur des continents séparés. Les deux genres principaux de Gammaprotéobactéries identifiées, *Erwinia* et *Serratia*, sont connus pour produire des métabolites antibiotiques et elles auraient un rôle de défense chimique identique à celui des actinobactéries, par ailleurs peu abondantes dans les échantillons analysés (Seipke *et al.* 2013).

De manière générale, les études du microbiote bactérien cuticulaire chez des fourmis sont majoritairement centrées sur la famille des attines et sur l'isolement d'actinobactéries dont le rôle principal dans ces interactions est la défense de l'hôte et de leur cultivar d'origine fongique. Un seul exemple à notre connaissance suggère une fonction de communication des métabolites du microbiote bactérien cuticulaire chez des fourmis *Pogonomyrmex barbatus*. Les individus ayant subi un traitement antibiotique de surface sont moins reconnus par leurs congénères non traités. Cette expérience semble indiquer que le microbiote bactérien pourrait être impliqué dans les mécanismes de communication intra spécifique au sein d'une colonie (Dosmann *et al.* 2016).

D. Assemblages de ces microbiotes bénéfiques

Il est admis que les microbiotes des organismes hôtes sont constitués de communautés microbiennes réparties de manière non aléatoire et que la probabilité de trouver les micro-organismes bénéfiques en plus grande abondance est plus importante par rapport à une situation où seul le hasard serait le facteur explicatif de la distribution des communautés microbiennes (Scheuring & Yu 2012). Cette hypothèse est supportée par le succès écologique des insectes sociaux qui ont mis en place de nombreux systèmes de défense chimiques et biologiques pour lutter contre la pression exercée par les souches pathogènes. Leurs présences dans différentes niches écologiques suggère l'utilisation de mécanismes adaptatifs pour recruter dans l'environnement des souches bactériennes bénéfiques productrices d'une grande diversité chimique de métabolites antimicrobiens (Kaltenpoth 2009). Chez les fourmis attines, une hypothèse propose que l'hôte favorise d'une part la présence d'actinobactérie sur la cuticule et d'autre part la compétition des communautés bactériennes pour la production des molécules antimicrobiennes (Sen *et al.* 2009). Ainsi il y aurait un équilibre entre une transmission verticale des micro-organismes bénéfiques et le recrutement de bactéries environnementales plus performantes.

IV. Les métabolites de la cuticule des fourmis

Les métabolites cuticulaires des fourmis appartiennent à différentes familles chimiques et sont essentiellement produits d'une part par les nombreuses glandes excrétales des produits naturels à la surface de la cuticule et d'autre part par les micro-organismes cuticulaires *in situ*, bien que ce deuxième point soit peu étudié. La diversité chimique des métabolites cuticulaires des fourmis est présentée dans les parties suivantes.

A. Les hydrocarbures cuticulaires

Les hydrocarbures cuticulaires, « cuticular hydrocarbons (CHCs) », sont des chaînes carbonées saturées ou insaturées, c'est à dire possédant des doubles liaisons, qui peuvent être méthylées. Plus de 1000 CHCs ont été décrits chez les fourmis et ils sont classés en 10 groupes principaux (Table 1). Par ordre d'occurrence décroissant, ces groupes sont : n-alcanes > monométhylalcanes > diméthylalcanes > alcènes > diènes > triméthylalcanes > méthylalcènes > méthylalcadiènes > triènes > tétraméthylalcanes (Martin & Drijfhout 2009). Les CHCs sont biosynthétisés par des cellules spécialisées associées aux cellules épidermiques, les oenocytes. Leur biosynthèse est similaire à celle des acides gras chez les vertébrés par assemblage de plusieurs unités malonyl-CoA en présence de l'enzyme ACP (3-ketoacyl carrier protein). Les réactions de méthylation et de déshydrogénation sont ensuite mises en jeu (Juárez *et al.* 1992). Les CHCs nouvellement synthétisés sont pris en charge par des protéines de transport, les lipophorines, qui assurent leur accumulation sur l'épicuticule. Les mécanismes impliqués dans leur transport ne sont pas tous connus (Blomquist & Bagnères 2010). Notons que la configuration absolue des carbones asymétriques portant un méthyle dans les structures des CHCs s'est avérée être systématiquement de type R suggérant ainsi que la voie de biosynthèse des réactions de méthylation est particulièrement conservée chez les insectes (Bello *et al.* 2015).

Les CHCs assurent deux rôles principaux. Le premier est d'éviter la dessiccation des insectes en formant une couche hydrophobe qui limite significativement l'évaporation de l'eau présente dans l'organisme. Le deuxième rôle démontré est une fonction de communication portée par le signal chimique du mélange de CHCs. La reconnaissance intra- et interspécifique chez les insectes sociaux dépend entièrement du profil chimique des CHCs cuticulaires. Chez les insectes sociaux (guêpes, fourmis et abeilles) les principaux groupes de CHCs (n-alcanes, alcènes et méthylalcanes) sont présents très tôt dans l'histoire évolutive des Hyménoptères (van Wilgenburg *et al.* 2011) et l'évolution des structures chimiques suggère que la complexité du signal se caractérise par un nombre plus important de groupements méthyles et de doubles liaisons. Par exemple seul un nombre limité d'espèce possède des CHCs comportant à la fois ces deux fonctions organiques (Table 1).

Table 1 : L'occurrence des classes de CHCs (en pourcentage) dans 4 groupes d'Hyménoptères. Le nombre d'espèces considérées dans chaque groupe (N) est indiqué (van Wilgenburg, Symonds, et Elgar 2011).

Chemical Class		Symphyta N=2	Parasitica N=27	Aculeate Wasps N=39	Ants N=95	Bees N=78
<i>n</i> -Alkanes		100	100	100	99	100
<u>Olefins:</u>						
Alkenes		100	63	82	78	91
Alkadienes		100	19	13	29	35
Alkatrienes					1	
Alkatetraenes			4			
<u>Methylalkanes:</u>						
Monomethylalkanes		50	78	85	98	35
Dimethylalkanes		50	70	62	83	4
Trimethylalkanes			30	8	23	1
Tetramethylalkanes			30		1	
<u>Methylolefins:</u>						
Methylalkenes				3	3	3
Methylalkadienes					3	

B. Les glandes exocrines : cas de la glande métapleurale

Chez les hyménoptères la glande métapleurale est exclusivement présente chez les fourmis et son origine phylogénétique date du Crétacé (Hölldobler & Wilson 1990). Cette glande située sur la face postérieure du mesosoma chez les fourmis excrète en continue des produits naturels sur la cuticule via un orifice ouvert (Figure 9). Elle est souvent absente chez les mâles et chez les familles Formicinae et Myrmicinae (Yek & Mueller 2011). Plusieurs hypothèses ont été émises concernant la fonction de cette glande. La plus probable est la fonction d'hygiène associée aux propriétés antiseptiques des métabolites sécrétés. Chez les espèces pourvues d'une glande métapleurale, les fourmis touchent l'orifice de cette glande avec leurs membres et se recouvrent le corps de la sécrétion produite. Il a été montré que ce phénomène est accru lors d'une infection expérimentale (Fernández-Marín *et al.* 2006). Il est suggéré que l'acidité de ces sécrétions ou/et la présence de molécules antibiotiques inhibent le développement des micro-organismes sur la cuticule. Le pH des sécrétions de la glande métapleurale est de 2,5 chez les *Atta sexdens*, de 3,0 à 3,5 chez *Myrmica rubra* et comprise entre 3 et 4 chez le genre *Crematogaster* (Cr.) et chez plusieurs fourmis de la sous famille des Ponerines (Maschwitz *et al.* 1970; Maschwitz). L'hypothèse selon laquelle l'acidité des sécrétions de la glande métapleurale est positivement corrélée à l'activité antibactérienne est

contestée car d'autres glandes exocrines (glande à poison) possédant également des sécrétions acides ne présentent pas d'activités antibactériennes avérées (Yek & Mueller 2011).

Les composés sécrétés par la glande métapleurale ont été analysés pour seulement quelques espèces modèles en myrmécologie. Ainsi 43, 19 et 16 composés ont respectivement été identifiés chez les fourmis coupeuses de feuille (*Atta* sp. *Acromyrmex* sp.), les fourmis de feu (*Solenopsis* sp.) et les *Cr.* sp. Les acides carboxyliques et les acides gras de longueur de chaîne variable dominant chez les fourmis coupeuses de feuille et chez les fourmis de feu (*Solenopsis* sp.), alors que les composés phénoliques dominant chez les *Cr.* sp (Yek & Mueller 2011). La présence de composés protéiques a également été décrite, ce qui corrobore avec la présence de réticulum endoplasmique dans la glande métapleurale qui atteste de la biosynthèse de peptides ou protéines (Schoeters, E. & Billen, J. 1991).

D'autres glandes exocrines moins étudiées, comme les glandes mandibulaires par exemple, contribuent également à la défense et/ou à la protection antiseptique (Schoeters & Billen).

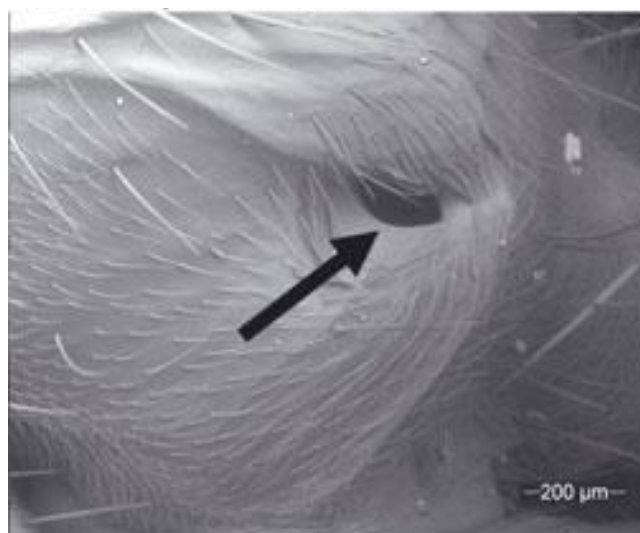


Figure 9 : Ouverture de la glande métapleurale chez *Paraponera clavata* (Yek & Mueller 2011).

C. Les métabolites des micro-organismes cuticulaires

La cuticule des fourmis est colonisée par des communautés bactériennes productrices de métabolites qui se retrouvent sur la cuticule et se mélange avec les métabolites produits par les fourmis. Les travaux antérieurs se sont principalement intéressés à identifier ces métabolites d'origine bactérienne à partir de la culture de souches isolées d'actinobactéries cuticulaires. Le métabolome bactérien cuticulaire produit *in vivo* reste difficilement accessible et inconnu. Un seul exemple à notre connaissance décrit *in vivo* un métabolite cuticulaire d'origine bactérienne

en utilisant comme technique analytique l'imagerie par spectrométrie de masse. Cette technique d'analyse permet de visualiser la distribution spatiale d'une entité moléculaire à la surface ou dans un espace en 3D au sein d'un organisme biologique. A titre d'exemple, l'imagerie par spectrométrie de masse a été utilisée pour visualiser la distribution spatiale des métabolites microbiens directement à partir de culture bactérienne sur milieu solide (Stasulli & Shank 2016). Ainsi, la distribution spatiale de la valinomycine (Figure 11) présente sur la cuticule de la fourmi *Acromyrmex echiniator* a été observée par imagerie MALDI (matrix-assisted laser desorption ionization) (Figure 10). Cet antibiotique est produit par une actinobactérie cuticulaire du genre *Streptomyces* (Schoenian *et al.* 2011) et serait impliquée dans la défense chimique du champignon cultivé par ces fourmis champignonnistes contre l'invasion de souches fongiques pathogènes.

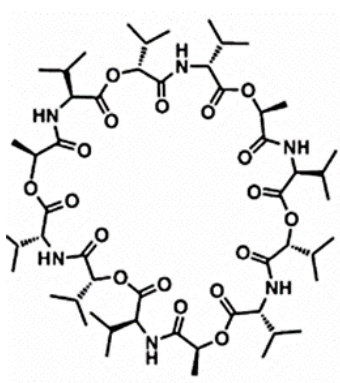


Figure 10 : Structure de la valinomycine isolée du *Streptomyces* associé à *Acromyrmex echiniator*.

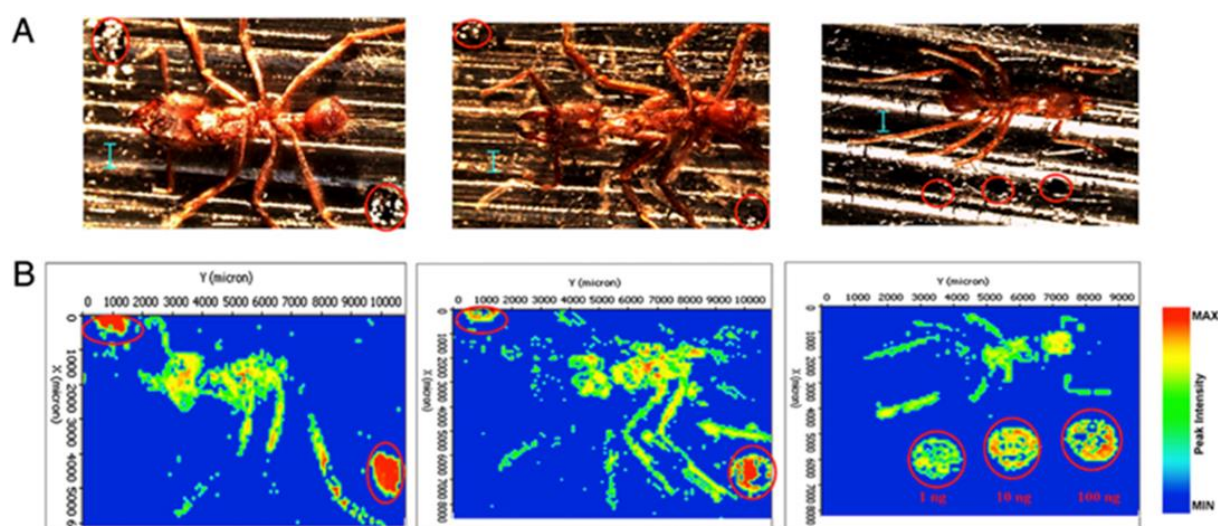


Figure 11 : Images microscopiques (A) et images MALDI (B) d'*Acromyrmex echiniator*. Des spots de valinomycine (rouge) sont déposés à côté de la fourmi et permettent d'estimer la concentration (Schoenian *et al.* 2011).

V. Les jardins de fourmis : un cas de parabiosis

Les jardins de fourmis ont été identifiés pour la première fois par Ernest Ule en 1901 au Brésil (Ule E 1901) et ils sont composés d'une plante épiphyte abritant une colonie de fourmis dans ses racines. Le jardin de fourmi est l'ensemble composé par la plante et la colonie de fourmis. Les fourmis intègrent les graines de certaines espèces d'épiphytes dans les parois du nid en carton qu'elles construisent dans les arbres, puis les épiphytes se développent et forment la structure du nid (Orivel & Leroy 2011). Seulement quelques espèces de fourmis sont connues pour former des jardins de fourmis. En effet elles doivent être capables de construire un nid en carton arboricole riche en humus et d'incorporer des graines pour lesquelles elles présentent une attraction particulière dans les parois du nid. Dans les zones néotropicales, une dizaine d'espèces sont décrites dans les jardins de fourmis, mais les plus abondantes sont *Ca. femoratus* (taille des ouvrières comprise entre 4 à 9 mm) et *Cr. levior* (taille des ouvrières comprise entre 2 à 3 mm) (Corbara *et al.* 1999)(Figure 12). Les espèces de plantes épiphytes les plus abondantes sont la broméliacée *Aechmea mertensii*, l'aracée *Anthurium gracile* et la gesneriacée *Codonanthe crassifolia* (Corbara *et al.* 1999) (Figure 13).



Figure 12 : Photographie de *Camponotus femoratus* (à droite) et de *Crematogaster levior* (à gauche) (auteur : Alain Dejean).

Camponotus femoratus et *Cr levior* ont une interaction biologique étroite appelée parabiose. Cette association est non obligatoire; les colonies des deux espèces se tolèrent, partagent le même nid, exploitent les mêmes ressources mais gardent leurs couvains séparés (Forel A 1898; Wheeler WM 1921). Ce type d'association est rare et n'a été décrit que pour quelques paires d'espèces dans le monde. Les apports des deux espèces à la plante hôte sont bénéfiques pour la plante, notamment parce qu'ils enrichissent l'environnement en minéraux et en matières organiques. L'humus des jardins de fourmis est plus riche en matières organiques que le sol proche (Belin-Depoux M 1991). Au-delà de l'apport nutritif, les épiphytes bénéficient

de la protection des fourmis et de la dispersion des graines. Les fourmis bénéficient d'un site de nidification adapté à leur physiologie, à leur système social, et se trouvent à proximité des ressources nutritives nécessaires à leur développement comme par exemple les fruits et les nectars extra-floraux et foliaires des plantes hôtes.



Figure 13 : Deux stades de développement des jardins de fourmis avec *Camponotus femoratus* et *Crematogaster levior*, mature à gauche et jeune à droite avec les épiphytes *Codonanthe calcarata* et *Aechmea mertensii* (Orivel & Leroy 2011).

Une question demeure concernant les mécanismes régissant la parabiose : comment ces deux espèces de fourmis se tolèrent-elles dans le même nid ? L'une des hypothèses suggérées est le mimétisme chimique. Des travaux récents menés en Guyane française sur la diversité des CHCs des deux espèces montrent qu'il existe deux chémotypes chez *Cr. levior* et les deux populations vivent séparément dans les jardins de fourmis avec *Ca. femoratus*. Étonnamment les deux chémotypes des CHCs présentent peu de recouvrement avec les CHCs de *Ca. femoratus*, d'après les analyses par GC/MS, ce qui suggère que le mimétisme chimique n'explique pas la tolérance de ces espèces entre elles (Emery & Tsutsui 2013, 2016). Une autre hypothèse propose la production de molécules apaisantes limitant le comportement agressif de l'espèce de taille la plus grande, *Ca. femoratus*. L'analyse de la composition des métabolites cuticulaires a permis d'isoler sur la cuticule de *Cr. modiglianii* un composé décrit pour la première fois, nommé crématoénone (Figure 14), qui réduit de façon significative le comportement agressif de *Ca. rufifemur* (Menzel *et al.* 2013).

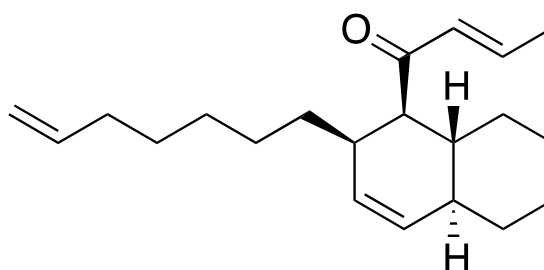


Figure 14 : Structure chimique de la crématoénone (Menzel *et al.* 2013).

Etat de l'art général des méthodes

I. Méthodes d'évaluation de l'activité antibiotique

A. Présentation et contexte

Les chimistes des produits naturels s'intéressent avant tout à la présence de molécules antibiotiques. La première étape du long parcours de l'identification d'une molécule active est la mise en culture liquide d'une souche d'intérêt pour extraire les molécules biosynthétisées dans le milieu de culture et obtenir un extrait (mélange de plusieurs molécules). Cet extrait fera l'objet d'un test pour évaluer sa capacité à inhiber le développement d'un pathogène. Une multitude de tests permettent d'évaluer le pouvoir d'inhibition d'une souche, d'un extrait ou d'une molécule pure. Les études d'écologie décrivant la diversité des souches cultivables d'un insecte ou d'une plante vont préférentiellement utiliser des tests d'antagonisme : cela consiste à mettre en culture la souche d'intérêt avec une souche pathogène et de déterminer s'il existe une inhibition de la croissance du pathogène. La coculture permet en effet de mimer les interactions entre micro-organismes dans la nature et donc d'induire la production de métabolites induits par le stress provoqué par la présence de la souche pathogène.

Lors d'un test sur milieu solide d'un extrait, le principe est également d'observer une zone d'inhibition provoquée par la présence de l'extrait alors que lors d'un test en milieu liquide, c'est l'observation de la concentration bactérienne via la lecture de la densité optique (DO) contre celle d'un témoin qui permettra de conclure quant au pouvoir d'inhibition de l'extrait. L'approche est identique pour une molécule pure testée à différentes concentrations permettant de déterminer une Concentration Inhibitrice à 50% (CI_{50}) correspondant à la concentration qui inhibe 50% de la croissance de la souche testée.

Deux grands types de méthodes existent : les méthodes dites de « diffusion » où l'on observe si une diffusion des métabolites dans le milieu inhibe la croissance d'une souche test et les méthodes dites de « dilution » où l'on teste les métabolites à différentes concentrations pour connaître la CMI (Concentration Minimum Inhibitrice) de ces métabolites. Une liste non exhaustive de ces méthodes est résumée dans la (Table 2) :

Table 2 : Description des méthodes d'évaluation de l'activité antibiotique.

Méthode	Micro-organisme testé	Principe	Application	Avantage	Inconvénient
Diffusion	Agar disk-diffusion	Une boîte de Petri est inoculée avec la souche testée. Un disque de papier filtre imprégné du composé à tester est placé au centre de la boîte. Le composé diffuse dans l'agar et inhibe le développement de la souche testée. Le diamètre de la zone d'inhibition est mesuré.	Tests de la susceptibilité des souches pathogènes aux antibiotiques = antibiogramme	Standardisée (CLSI Clinical and Laboratory Standards Institute), aide à la décision pour la prescription d'antibiotiques à l'hôpital, faible coût, simplicité	Ne détermine pas les effets bactéricides, bactériostatiques et la concentration minimale inhibitrice (CMI) du composé.
	Antimicrobial gradient	Une boîte de Petri est inoculée avec la souche testée. Une bande de papier filtre imprégné d'un gradient du composé à tester est placée au centre de la boîte. Le composé diffuse dans l'agar et inhibe le développement de la souche testée. La CMI est déterminée à l'intersection de la bande et de l'ellipse de l'inhibition.	Détermination de la CMI des composés	Bonne corrélation entre la CMI obtenue et celle obtenue lors des méthodes de dilutions	Coûteux
	Agar well diffusion	Une boîte de Petri est inoculée avec la souche testée. Un trou est réalisé de manière stérile au centre de la boîte où l'extrait à tester est déposé et diffuse. Le diamètre de la zone d'inhibition est mesuré.	Evaluer l'activité antimicrobienne des extraits de plantes ou de micro-organismes.	Similaire à Agar disk-diffusion	Non standardisée
	Agar plug diffusion	Une boîte de Petri est inoculée avec la souche testée. La souche d'intérêt est cultivée sur le milieu approprié permettant ainsi la production des métabolites directement dans l'agar. Un morceau d'agar (agar plug) est découpé et déposé au centre de la boîte inoculée. Le diamètre de la zone d'inhibition est mesuré.	Mettre en évidence l'antagonisme entre micro-organismes	Similaire à Agar disk-diffusion	Non standardisée
	Cross streak	La souche d'intérêt est inoculée sur une bande au centre de la boîte de Petri permettant ainsi la production des métabolites directement dans l'agar. Après un temps d'incubation, la souche à tester est inoculée sur une bande perpendiculaire à la bande centrale. Après incubation la zone d'inhibition est mesurée.	Mettre en évidence l'antagonisme entre micro-organismes, utilisée dans le cadre de screening	Plus rapide que agar plug diffusion, forte croissance d'inhibition	Non standardisée
Poisoned food	Champignons	Le composé antifongique testé est incorporé au milieu et coulé dans une boîte de Petri. Un morceau d'agar rond sur lequel se développe la souche testée est déposé au centre de la boîte. Après incubation le diamètre de la croissance du champignon est mesuré.	Evaluation de l'activité antifongique		Non standardisée
Dilution	Thin layer chromatography : Direct bioautography	L'extrait étudié est élué sur une plaque de chromatographie sur couche mince. La suspension de la souche testée est pulvérisée sur la plaque. Après un temps d'incubation, des révélateurs à base de sels de tétrazolum sont utilisés pour observer les tâches où il n'y a pas de croissance.	Identifier les composés antibactériens et antifongiques	Technique plus simple pour détecter des substances antifongiques	
	Broth dilution	Les composés sont testés à différentes dilutions dans les puits d'une plaque 96 puits où une suspension de la souche à tester est ajoutée. La lecture de la croissance de la souche testée est réalisée à l'œil, en mesurant une DO (densité optique) ou avec des techniques de coloration.	Détermination de la CMI des composés	Standardisée (CLSI), rapidité, peu coûteux, miniaturisation et reproductibilité	
	Agar dilution	Les composés sont intégrés dans le milieu agar et testés à différentes dilutions. Puis les boîtes de Petri sont inocuées avec une suspension de la souche à tester. La lecture de la croissance de la souche testée est réalisée à l'œil et la CMI est la boîte avec la plus petite concentration permettant l'inhibition totale de la souche testée.	Détermination de la CMI des composés	Permet de tester des composés colorés (non testés avec broth dilution)	
	Time-kill test (time-kill curve)	3 tubes sont réalisés avec une suspension de la souche à tester : 2 tubes ont des concentrations différentes du composé à tester et le troisième tube est le contrôle. Pendant l'incubation, des comptages des cellules vivantes sont réalisés à plusieurs intervalles.	Détermination de l'effet bactéricide et fongicide	Standardisée (CLSI)	
	ATP bioluminescence assay	Mesure de la concentration en ATP produit par la souche à tester.	Estimer la population microbienne pour un échantillon	Rapide, test antimicrobien in vivo ou in situ	

B. La méthode cross-streak

La cross-streak est une méthode d'évaluation de l'activité d'antagonisme d'une souche contre des souches tests pathogènes. L'équipe du professeur Waksman est à l'origine en 1945 de la mise au point de cette méthode. Cette technique facile à mettre en place a été reprise par de nombreux laboratoires. En 1947, cette méthode est utilisée pour tester l'activité d'antagonisme d'actinobactéries contre le bacille de la tuberculose (Williston *et al.* 1947). De nombreuses études d'écologie microbienne décrivant la diversité cultivable des actinobactéries du sol marin (Valli *et al.* 2012; Claverías *et al.* 2015), des intestins de poisson (Jami *et al.* 2015) ou des nids de guêpes en carton (Madden *et al.* 2013) ont testés l'antagonisme de leurs souches isolées afin d'argumenter en faveur de l'hypothèse du rôle défensif de ces actinobactéries pour leur hôte.

La méthode cross-streak a récemment permis d'identifier que les actinobactéries de la cuticule d'attines (*Acromyrmex sp.*) inhibent le développement des souches fongiques isolées à partir de la cuticule des mêmes individus. Dans ce cas, l'activité antimicrobienne des actinobactéries n'est pas spécifique du champignon *Escovopsis* pathogène de leurs jardins et s'étend à d'autres souches fongiques. Ce test est simple dans sa mise en œuvre et permet d'évaluer le rôle biologique des actinobactéries dans cette interaction complexe (Dângelo *et al.* 2016). Un autre type de test d'antagonisme a été proposé pour tester l'antagonisme des souches d'actinobactéries associées à la cuticule de fourmis champignonnistes. Une des deux souches est inoculée au centre d'une boîte de pétri « *resident strain* » et se développe pendant une semaine avant que la deuxième souche « *intruder strain* » soit introduite tout autour de la première. La zone d'inhibition en forme de cercle entre les deux est mesurée. Des relations d'antagonismes entre souches présument du rôle préventif de ces antagonismes pour empêcher l'invasion de la cuticule par de nouvelles souches (Poulsen *et al.* 2007; Van Arnam *et al.* 2015).

II. Techniques de chimie analytique

A. La chromatographie

La chromatographie en phase liquide à polarité de phase inverse est un outil incontournable d'analyse de mélanges complexes de molécules solubles. Les métabolites solubilisés dans un solvant organique sont injectés dans une phase mobile qui traverse une colonne comportant une phase stationnaire solide apolaire, composée le plus souvent de chaînes alcanes (C18). Cette technique de séparation chromatographique permet la séparation des molécules polaires de haut poids moléculaire tout en limitant leur dégradation ou leur décomposition. Avec les sources d'ionisation modernes à pression atmosphérique, la spectrométrie de masse peut être directement couplée en sortie de colonne chromatographique

et utilisée comme détecteur, on parle ainsi de couplage LC-MS (*Liquid chromatography Mass Spectroscopy*).

Les systèmes de chromatographie liquide de dernière génération développés par le constructeur Waters (*Ultra Performant Liquid Chromatography*, UPLC) permettent de travailler à très haute pression (jusqu'à 1000 bars) avec des colonnes dont la granulométrie est inférieure à 2 μm . Ceci permet d'augmenter la vitesse d'élution de la phase mobile et donc d'améliorer significativement la séparation des molécules. Les pics sont détectés avec une plus grande résolution et un rapport signal sur bruit plus élevé (Nordström *et al.* 2006). Le temps d'élution en UPLC est largement inférieur à celui obtenu en HPLC (*High-Performance Liquid Chromatography*) ce qui réduit la consommation des solvants de la phase mobile et le coût de l'analyse.

La chromatographie en phase gazeuse (*Gaz chromatography*, GC) est une technique différente de la chromatographie en phase liquide et fait l'objet d'une description ci-dessous car c'est la technique la plus utilisée en écologie chimique, pour l'analyse des hydrocarbures cuticulaires des fourmis par exemple. Brièvement, cette technique nécessite dans un premier temps de solubiliser les métabolites dans un solvant organique avant de les introduire dans l'injecteur qui vaporise les molécules à une température comprise entre 200°C et 300°C. Les molécules en phase gazeuse sont alors séparées dans une colonne chromatographique selon leur affinité avec la phase solide de la colonne avant d'être analysées par le détecteur, un spectromètre de masse le plus souvent (*Gaz chromatography Mass spectroscopy*, GC-MS). Ainsi, la GC-MS ne permet pas l'analyse des molécules non volatiles, soit parce qu'elles ont un haut poids moléculaire (≥ 1200 daltons), soit parce qu'elles ont un point d'ébullition supérieur à 300°C (Burgot & Burgot 2011).

B. La spectrométrie de masse

La spectrométrie de masse est une technique de détection de molécules sous forme ionisée permettant de déterminer leur masse moléculaire ou celle de leurs fragments. Elle peut être appliquée après une séparation chromatographique par exemple. C'est la mesure du rapport masse-sur-nombre de charges communément noté m/z de molécules ionisées en phase gazeuse dans la source du spectromètre de masse. Un spectromètre de masse se compose de trois éléments : une source, un analyseur et un détecteur. La source assure l'ionisation et le passage à l'état gazeux des molécules, l'analyseur sépare les molécules ionisées en fonction de leur valeur m/z (masse/charge) et le détecteur enregistre le signal correspondant aux différentes molécules et/ou fragments en recevant un flux d'ions et émet un signal proportionnel au courant ionique généré. Parmi les différentes technologies en spectrométrie de masse pour ioniser les

molécules, l'électronébulisation (*electrospray ionization*) est une méthode d'ionisation développée dans les années 1970 et considérée comme douce (Clegg & Dole 1971). Elle consiste à soumettre un champ électrique à une dispersion d'un liquide sous forme de gouttelettes à pression atmosphérique. Remarquons que les hydrocarbures cuticulaires ne sont pas ionisés dans ces conditions douces et ils ne sont donc pas détectés par cette technique d'analyse.

Il est possible d'associer en tandem plusieurs types d'analyseur dans un spectromètre de masse et de combiner leurs points forts. Ces appareils appelés Q-TOF sont constitués d'un analyseur primaire, le double quadripôle, et d'un analyseur secondaire à temps de vol (*Time of Flight*, TOF). Le quadripôle procure une grande efficacité tandis que le TOF apporte une excellente sensibilité. Cette combinaison d'analyseurs permet d'obtenir des résolutions et des précisions de masse des ions produits bien supérieures à une configuration de type triple quadripôle.

L'analyse métabolomique des cuticules de fourmis effectuée lors de ce travail de thèse utilise pour la première fois la chromatographie UPLC couplée à un spectromètre de masse de haute résolution (*High Resolution Mass Spectroscopy*) (UPLC-HRMS). Dans le cadre des travaux présentés ici, le spectromètre de masse SYNAPT G2 HDMS de la société Waters a été utilisé (Figure 15). Il s'agit d'un instrument hybride Quadrupole Time of Flight QTOF à haute résolution composé d'une source d'ionisation par électronébulisation, d'un quadripôle, de deux cellules de collision, d'une cellule de mobilité ionique, d'un analyseur TOF à injection orthogonale et d'un détecteur.

Il est important de souligner que dans le cadre de cette thèse nous nous intéressons aux métabolites cuticulaires d'origine bactérienne. Par conséquent l'utilisation de ce mode d'ionisation douce (ESI) nous permettra de détecter principalement les molécules ionisables produites par les fourmis et les produits naturels d'origine microbienne recherchés. Les CHCs sont des chaînes hydrocarbonées qui ne sont pas ionisables dans ces conditions d'ionisation et sont donc exclus de l'analyse. Ceci permettra d'améliorer la détection des métabolites microbiens d'intérêt. Ce travail représente le premier exemple en chimie analytique de l'analyse métabolomique des cuticules d'insectes par UPLC/MS.

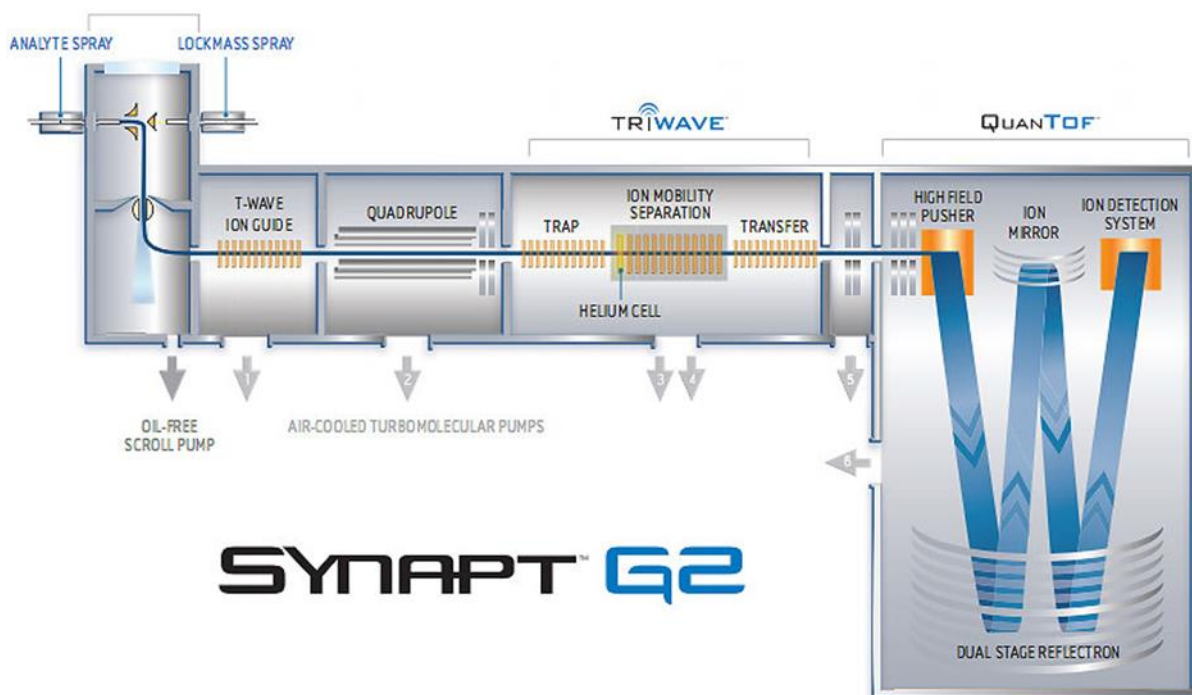


Figure 15 : Schéma du spectromètre de masse SYNAPT G2 HDMS (Waters Corporation, Manchester, UK).

Les données sont acquises en mode centroïde et se présentent sous la forme d'un chromatogramme pour chaque échantillon. Le mode centroïde est le traitement du signal du détecteur par un convertisseur, qui remplace les pics par un trait qui est le centre de gravité du pic pour un ion dont on calcule ensuite le rapport m/z . Sur chaque chromatogramme est représentée l'intensité du courant ionique total (*Total Ion Current, TIC*) en fonction du temps de rétention. Pour chaque temps de rétention, il existe une série d'ions détectés ayant des rapports m/z différents. Ainsi, chaque ion détecté possède trois caractéristiques : son rapport m/z , son temps de rétention et l'intensité de son signal (Figure 16). La combinaison entre le temps de rétention d'un ion et sa valeur m/z permet de le caractériser et constitue une variable appelée ion dans la suite du manuscrit.

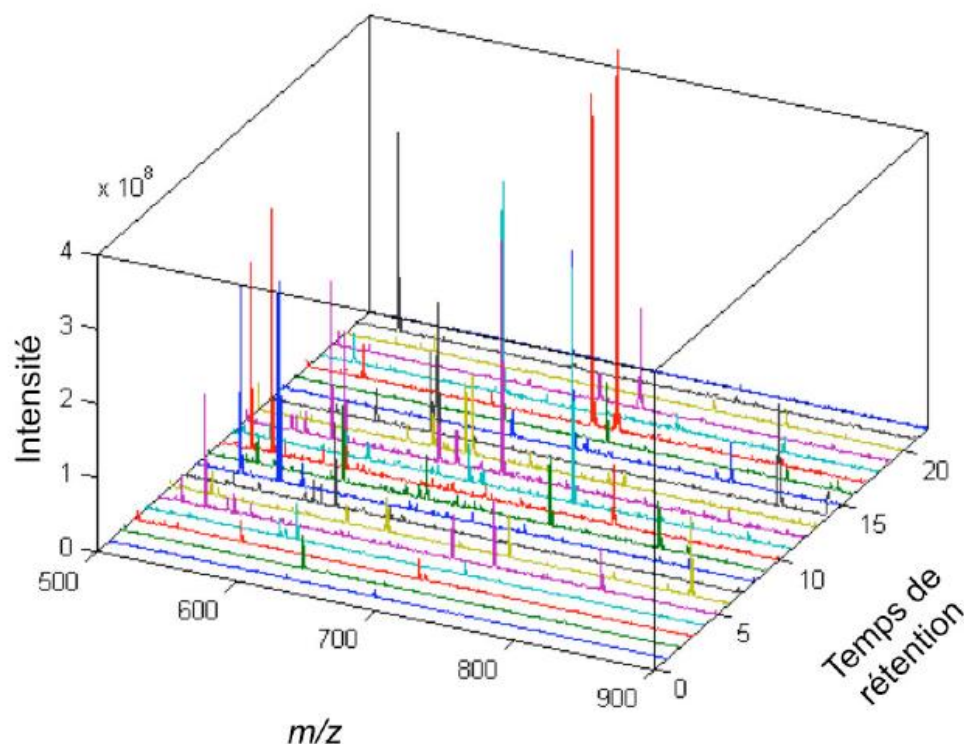


Figure 16 : Représentation tridimensionnelle d'un chromatogramme obtenu par *Liquid Chromatography-Mass Spectrometry*.

C. Apports et principe des réseaux moléculaires

Les dernières avancées en bioinformatique et en chimie analytique, en particulier en spectrométrie de masse, ont totalement révolutionné la découverte et l'analyse des produits naturels (Medema & Fischbach 2015). Récemment, les réseaux moléculaires (*molecular networking*) nouvellement développés constituent une approche prometteuse pour interpréter les données MS/MS et automatiser la recherche dans les bases de données pour l'identification rapides des métabolites dans des mélanges complexes (Yang *et al.* 2013). Un avantage majeur dans l'utilisation des réseaux moléculaires pour effectuer le travail de déréplication est la possibilité d'identifier de façon simple les composés connus et de les discriminer des analogues potentiels. Cette approche facilite considérablement l'analyse chimique d'un échantillon, composé par une grande diversité de métabolites inconnus, qui peuvent être reliés aux familles chimiques déjà connues. Par analogie avec l'annotation des taxons à partir de séquences ADN en métabarcoding, la principale limitation des réseaux moléculaires est la disponibilité et la qualité des bases de données spectrales MS/MS. La plateforme en ligne gratuite *Global Natural Product Social molecular networking* (GNPS) a été développée par l'équipe de Pieter Dorrestein

et Nuno Bandeira (<https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp>) et elle est dédiée à l'interprétation et l'indexation des données MS² (Wang *et al.* 2016).

Les réseaux moléculaires nécessitent l'utilisation de données acquises en MS². Lors d'une analyse MS², le spectromètre de masse en tandem sélectionne un ion précurseur avec le premier analyseur puis cet ion est fragmenté, et les fragments ainsi formés sont analysés afin d'obtenir un spectre MS². Ce dernier est composé uniquement d'ions produits par la fragmentation de l'ion précurseur et le spectre obtenu est donc riche en informations structurales, contrairement aux spectres MS (sans fragmentation en ESI). Les molécules similaires partagent des voies de fragmentation similaires conduisant à des fragments ou des pertes de neutre en commun, ce qui est la base de la construction des réseaux moléculaires. Ainsi, les spectres MS² d'un jeu de données sont comparés deux à deux par alignement spectral, puis un indice de corrélation, appelé similarité cosinus (*cosine score*, CS), est attribué entre chaque paire de spectres. Les spectres ayant un CS supérieur à une certaine valeur (typiquement 0,7) ont une forte similarité spectrale ; ils sont alors considérés comme issus de la fragmentation de constituants chimiquement apparentés (Nothias-Scaglia *et al.* 2015). Ainsi, les ions précurseurs sont connectés grâce au CS ce qui regroupe les composés en famille chimique sur le réseau.

D. La métabolomique

1) Définitions

Le métabolome est l'ensemble des entités chimiques de faible poids moléculaires d'un échantillon biologique. Le métabolome est variable et évolue au cours du temps. D'un point de vu sémantique, le terme métabolome a été admis en 1998 (Oliver *et al.* 1998) et par analogie il fait référence aux termes génome, protéome et transcriptome. La détection et la quantification de molécules ionisées en sortie du spectromètre de masse, appelés plus simplement ions dans la suite de ce manuscrit (*features*) concerne un spectre large de familles chimiques détectables par le spectromètre comme les sucres, des lipides, les terpènes, les alcaloïdes, les composés aromatiques etc.

La métabolomique est la production et l'analyse statistique des données chimiques d'un ou plusieurs organismes biologiques. Ces données sont obtenues généralement par différentes techniques d'analyse telles que la spectroscopie par résonance magnétique nucléaire (RMN) ou par la spectrométrie de masse (MS). La métabolomique par spectrométrie de masse utilise principalement des systèmes chromatographiques couplés à un spectromètre de masse (GC/MS, HPLC/MS, UPLC/MS).

Par la détection directe et la quantification des métabolites, c'est l'outil qui permet d'étudier le produit final de l'ingénierie du vivant pour interagir avec son environnement. La métabolomique est ainsi plus proche de la métaprotéomique que du séquençage de l'ADN (métagénomique) ou de l'ARN (métatranscriptomique). Les métabolites peuvent être reliés à un groupe de gènes à l'aide des voies de biosynthèse connues entre les gènes et les enzymes codant pour ces métabolites. Les interactions entre les micro-organismes et leurs hôtes se font souvent par le biais d'échanges de métabolites, ce qui fait de la métabolomique un outil essentiel pour la compréhension des fonctions des communautés microbiennes (Franzosa *et al.* 2015).

2) *Prétraitement des données*

Des étapes de prétraitement des fichiers bruts de données sont effectuées afin d'extraire les variables d'intérêt du bruit et leurs informations pour l'interprétation. Ces étapes sont réalisées majoritairement avec le package [xcms](#) (Smith *et al.* 2006) de R et sont présentées brièvement ci-dessous :

- Conversion des fichiers : Les outils utilisés par la suite ne traitent pas les fichiers bruts constructeurs (.raw avec Waters) et nécessitent des formats libres. Le logiciel en ligne [Proteowizard](#) convertit les .raw en .mzXML (Kessner *et al.* 2008).
- Détection des pics : l'algorithme centWave (Tautenhahn *et al.* 2008) détecte les pics des données acquises en mode centroïde pour chaque échantillon.
- Appariement des pics : l'algorithme group.density permet d'apparier tous les pics correspondant à une même variable à travers les échantillons en fonction des temps de rétention et des valeurs m/z.
- Correction des temps de rétention : les groupes de pics précédents sont utilisés pour corriger une éventuelle dérive des temps de rétention au cours des injections avec l'algorithme obiwarp (Prince & Marcotte 2006). Les pics réalignés sont groupés une deuxième fois plus précisément.

Les données finales sont sous la forme d'une matrice de données brutes comprenant les individus, dont les QC (*Quality Control*) et les variables qui sont les ions.

3) *Contrôle de l'acquisition avec les QC*

Le bon déroulement de l'acquisition des données métabolomiques lors de l'analyse sur le spectromètre de masse est contrôlé via une analyse en composante principale (ACP) des données avec les QC (*Quality Control*) qui sont les contrôles qualité. Le QC est un échantillon réalisé à partir du mélange de tous les échantillons. Deux traitements des données brutes sont

nécessaires avant de réaliser cette analyse : la normalisation des données et le traitement des valeurs manquantes.

Des variations de la réponse mesurée non liées aux différences biologiques entre les individus sont observées lors d'analyses métabolomique comprenant de nombreux individus. Ces variations peuvent être liées à des petites différences de volumes lors de la préparation des *vials*, lors de l'injection des échantillons ou aux performances du spectromètre de masse. Les données sont normalisées pour corriger ces différences d'intensité de pics et pour stabiliser la variance dans la matrice de données. Ici une normalisation par somme est appliquée, chaque intensité est divisée par la somme totale des intensités d'un individu et multiplié par 100, la somme des intensités de tous les pics d'un individu est utilisée comme facteur de normalisation.

Tous les échantillons injectés (individus + QC) n'ont pas une valeur d'intensité pour tous les ions détectés de la matrice. Les valeurs manquantes pour un individu en métabolomique sont observées pour trois raisons : le métabolite n'est pas présent dans l'individu, le métabolite est présent mais à une concentration inférieure à la limite de détection de l'appareil et le métabolite est présent à une concentration supérieure à la limite de détection mais l'acquisition de cette donnée a échoué pour une raison technique. Tout en maintenant la structure des données, il est nécessaire de remplacer les valeurs manquantes par des valeurs différentes de 0 pour la robustesse des analyses multivariées (Gromski *et al.* 2014; Di Guida *et al.* 2016).

L'ACP est une méthode non supervisée exploratoire qui permet de déceler, lorsqu'elle est réalisée sur toute la matrice des données, si les QC sont bien regroupés. Les QC étant injectés tout au long de l'analyse, un alignement des QC sur la PCA signalerait d'un décalage anormal des temps de rétention par exemple.

III. Technique de génétique

A. Métabarcoding

Le challenge de l'identification morphologique des espèces est de les collecter puis de les classer pour les identifier. L'identification morphologique des espèces est ainsi un processus chronophage, coûteux et obligatoirement associé aux compétences de taxonomistes spécialisés. Le « barcoding moléculaire » ou « *DNA barcoding* » ou « code-barres génétique » est une technique récente de taxonomie moléculaire permettant la caractérisation génétique d'un individu ou d'un pool d'individus. Le métabarcoding désigne l'étude de tout un assemblage de populations à partir de l'ADN (i.e. un métagénome) contenu dans un échantillon environnemental (e.g. sol, sédiments, excréments). Il repose sur l'amplification par PCR d'une région commune à un groupe taxonomique donné qui est suffisamment variable pour pouvoir

différencier les espèces (Taberlet *et al.* 2012). Le métabarcoding est donc à différencier de la métagénomique car il se focalise sur une région génomique particulière alors que la métagénomique consiste à séquencer l'ensemble du métagénome. Ainsi, le métabarcoding permet d'identifier les taxons présents dans la communauté, alors que la métagénomique donne aussi accès aux gènes fonctionnels de la communauté contenue dans un échantillon environnemental, sous réserve que le séquençage du métagénome soit suffisamment profond (Riesenfeld *et al.* 2004). Dans le cadre du métabarcoding, il est possible d'utiliser des combinaisons d'amorces taguées différentes pour chaque échantillon pour l'étape d'amplification. Ceci permet de multiplexer plusieurs individus pour le séquençage et de retrouver à posteriori à quel individu les séquences appartiennent.

Pour les macro-organismes, les techniques d'identification des espèces par séquençage de l'ADN sont plus rapides mais dépendent de banques de données d'ADN de références. Concernant les bactéries, elles sont d'autant plus pertinentes que les techniques de culture standards permettent d'accéder à 1% ou moins de la diversité bactérienne dans la plupart des échantillons environnementaux (Amann *et al.* 1995). De plus, les caractères morphologiques microbiens sont largement insuffisants pour classer ces organismes, ne serait-ce qu'au niveau du phylum. Ainsi, l'introduction du séquençage des gènes codant pour les ARN ribosomiques en écologie microbienne a permis d'accéder à une diversité bactérienne jusqu'alors inconnue dans des systèmes variés tels, y compris à partir d'échantillons de fèces (Vo & Jedlicka 2014) ou stomacaux (Russell *et al.* 2009) ou de lessivas de cuticules d'insectes (Seipke *et al.* 2013).

B. Description de la technique de séquençage HTS

Le terme « séquençage à haut-débit » (HTS, *High-throughput sequencing*) regroupe l'ensemble des technologies de séquençage développées dans les années 2000. Aujourd'hui, on différencie les technologies à haut débit de 2nde génération qui permettent de séquencer des fragments courts, de celles dites de 3^{ème} génération permettant le séquençage d'une longue molécule d'ADN. Le marché des séquenceurs de 2nde génération est couvert par trois groupes que sont Roche, Illumina et Life Technologies ayant respectivement proposés de manière successive, leur première plateforme à savoir le 454, le Genome Analyser et enfin le SOLiD. Dans le cadre de ce travail de thèse, le séquençage a été réalisé sur les amplicons (résultats de l'amplification des barcodes par PCR) purifiés pour éliminer les nucléotides, amorces et autres « déchets » dans les échantillons. Les amplicons sont des fragments d'ADN courts (ca. 200 bp) qui peuvent être séquencés en pleine longueur sur la plateforme MiSeq (Illumina). C'est ce séquenceur de 2^{ème} génération et le protocole de lecture en sens et anti-sens 2*250 bp qui a été utilisé, et est illustré ci-dessous. Le séquençage a été réalisé à la Genotoul

(<http://www.genotoul.fr/>, Toulouse) en étroite collaboration avec le laboratoire Evolution et Diversité Biologique. A ce stade, le séquençage haut débit se déroule en 3 étapes :

1 : Préparation de la librairie

Les amplicons multiplexés sont soumis à une étape de ligation pendant laquelle des séquences adaptatrices particulières au séquenceur sont liguées aux amplicons. Le produit de ligation est ensuite étalé sur une microplaque contenant les mêmes adaptateurs greffés, auxquels les amplicons se fixent par complémentarité.

2 : Amplification

Elle s'effectue en émulsion (technologie ion torrent) ou sur support solide (technologie Illumina). Chaque fragment fixé sur le support grâce aux adaptateurs est amplifié environ 1000 fois pour former des « clusters » clonaux qui seront séquencés en parallèle. L'amplification permet ainsi, au bout d'un nombre de cycle suffisant, d'obtenir un nombre d'amplicons par cluster nécessaire au séquençage (Figure 17).

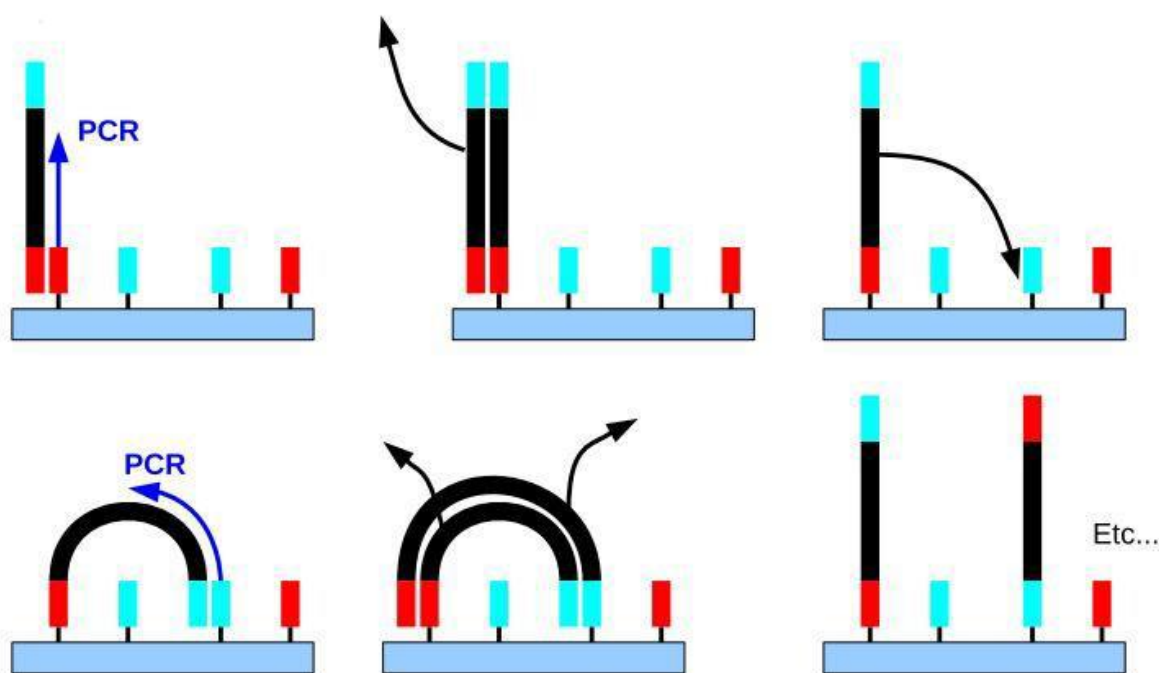


Figure 17 : Amplification en phase solide (HTS ; Illumina).

Le séquençage de l'ADN est réalisé par ajout de dideoxynucleotides marqués avec des fluorochromes. Ces dideoxynucleotides, substituts des nucléotides, bloquent l'ajout du nucléotide suivant. Les nucléotides sont donc ajoutés les uns après les autres et intercalés par des étapes de rinçages et de clivages. Ainsi lorsque le nucléotide complémentaire à la séquence

à amplifier se fixe, il empêche un autre de se fixer à la suite. L'étape de rinçage va permettre d'éliminer les nucléotides en surplus (non fixés) et de détecter les signaux fluorescents des nucléotides complémentaires fixés. Puis l'étape de clivage va enlever l'extrémité bloquante, permettant de passer au nucléotide suivant selon le même schéma.

3 : Séquençage Illumina

Il s'effectue grâce à la détection de photons. À chaque cycle, l'ajout d'un nucléotide se traduit par un signal fluorescent associé à chacun des quatre nucléotides de l'ADN. La lecture est possible car chaque nucléotide possède un marqueur fluorescent qui lui procure une couleur spécifique. Les amplicons ainsi séquencés, après plusieurs étapes de traitement bioinformatiques, sont ensuite comparés à des bases de références pour l'ADN bactérien, ce qui permettra d'assigner les lectures à des lignages bactériens

C. Les approches multi-omics :

Les techniques « omics » sont un ensemble d'outils et d'approches dont les quatre principaux sont la métagénomique, la métatranscriptomique, la métabolomique et la protéomique. La métagénomique est l'analyse de l'ensemble du matériel génétique d'un échantillon (i.e. du métagénome), la métatranscriptomique est l'analyse des molécules transcrites (i.e. ARN), la protéomique est l'analyse de l'ensemble des protéines et la métabolomique est l'analyse des petits composés chimiques présents dans l'échantillon. Chacune de ces techniques a permis, seule, à mieux comprendre de nombreux systèmes biologiques. Les combiner permet d'affiner cette compréhension. Par exemple, il est possible de caractériser une communauté bactérienne dans un substrat donné avec le métabarcoding, d'en comprendre les fonctions (métatranscriptomique et protéomique) et d'apprécier l'activité de la communauté avec la métabolomique (Figure 18). Cette approche multiple est appelée « multi-omics ». Deux grands types d'approches existent pour combiner l'information apportée par ces différentes méthodes : l'interprétation conjointe de données analysées séparément ou l'interprétation de données intégrées à une même analyse. Dans le premier cas, les composés importants de chaque type de donnée sont sélectionnés puis mis en relation via des analyses en réseau par exemple. Dans le cas de l'approche intégrée, des outils permettent de fusionner les données avant l'analyse des données et leur interprétation. Dans le cadre de ce travail de thèse, c'est la première approche qui sera proposée en perspective du chapitre 3. Une dernière approche consiste à travailler sur des modèles biologiques bien documentés avec une méthode d'intégration basée sur un modèle, où sont comparés des données expérimentales à des données prédites par le modèle (Beale *et al.* 2016).

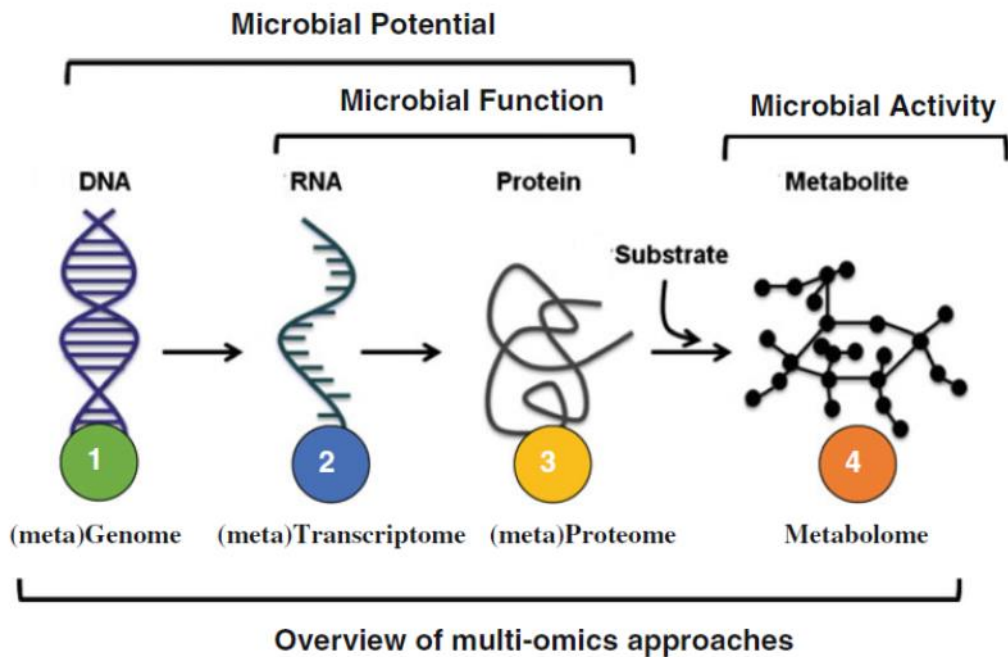


Figure 18 : L'intégration de données « multi-omics » pour faciliter l'analyse intégrée d'un microbiote (Beale *et al.* 2016).

De nombreux outils permettent aux chercheurs d'intégrer des données de plusieurs « omics », tout en ayant à disposition des outils statistiques pour normaliser les données par exemple. Le package R [mixOmics](#) utilisé dans le chapitre 3 (Lê Cao *et al.* 2009) regroupe un large éventail de méthodes multivariées permettant d'explorer et d'intégrer des données biologiques avec un accent particulier sur la sélection des variables. La sélection des variables est nécessaire pour réduire la taille des données omics où le nombre de variables explicatives et/ou de réponse (par exemple les gènes, les protéines, les métabolites) est nettement supérieur au nombre d'échantillons/individus (e.g. patients, cellules, souris). Dans ce contexte, les méthodes multivariées classiques ne sont pas adaptées aux analyses multivariées mais sont limitées pour les grands ensembles de données omics où le nombre de variables est beaucoup plus grand que le nombre d'échantillons.

Chapitre 1. Pouvoir antibiotique d'actinobactéries cuticulaires de fourmis.

I. Introduction

De nombreuses souches d'actinobactéries sont engagées dans des interactions mutualistes avec leur hôte où elles assureraient un rôle de défense qui améliore la survie de l'hôte. Dans les milieux terrestres, différentes souches d'actinobactéries ont été isolées sur la cuticule d'insectes, notamment chez les fourmis (Oh *et al.* 2009a), les coléoptères (Oh *et al.* 2009b) et les guêpes (Oh *et al.* 2011), où des molécules antibiotiques ont été identifiées à partir de la culture de ces souches, soutenant ainsi l'hypothèse de leur rôle défensif. L'étude des symbioses mutualistes entre les actinobactéries et les insectes sociaux représente par conséquent une source prometteuse pour la découverte de molécules bioactives originales (Kaltenpoth 2009). Les micro-organismes mutualistes (actinobactéries ou autres) à but défensif dans les interactions avec des hôtes eucaryotes constituent des candidats prometteurs pour la production d'antibiotiques à fort potentiel en santé humaine notamment car les composés de défense impliqués ne doivent pas nuire à l'hôte eucaryote et sont ainsi moins susceptibles de provoquer des effets secondaires sur l'homme.

En microbiologie, l'utilisation des cocultures, mélange de deux ou plusieurs micro-organismes, est une approche bio-inspirée qui vise à reproduire *in vitro* les interactions *in vivo* des communautés microbiennes à l'état naturel. Ces communautés interagissent entre elles par le biais de signaux chimiques de défense ou de communication. Les hypothèses proposées pour expliquer l'assemblage du microbiote cuticulaire bénéfique chez les fourmis renforcent la pertinence de la mise en place de cocultures. En effet la compétition intra-spécifique bactérienne induit la production de molécules d'intérêt (Bertrand *et al.* 2014). Donc, la stratégie de coculture de plusieurs souches bactériennes collectées sur un même hôte est pertinente. Cette approche permet d'augmenter l'expression de gènes codant pour des métabolites déjà produits en monoculture et permet l'expression de gènes restés silencieux en monoculture mais responsables de la formation de métabolites bioactifs originaux (Bertrand *et al.* 2014). Cette technique est intéressante plus particulièrement chez les actinobactéries où une forte portion de leur génome est considérée silencieuse, c'est-à-dire non codante (Doroghazi & Metcalf 2013; Tokovenko *et al.* 2016).

Afin de caractériser le potentiel bioactif du groupe des actinobactéries sur la cuticule des fourmis de Guyane, l'objectif de ce travail a été d'isoler diverses souches d'actinobactéries à partir de la cuticule des fourmis et d'identifier des métabolites secondaires aux propriétés

antibiotiques. L'échantillonnage constitué au fil des collectes était opportuniste et consistait à collecter un maximum d'espèces de fourmis aux modes de vie contrastés dans différents sites en Guyane (forêt primaire jusqu'aux abords des villes). Ceci dans le but d'accroître les possibilités d'isoler des souches d'actinobactéries originales. Les actinobactéries isolées lors de ce travail ont été identifiées par séquençage Sanger du gène codant pour l'ARNr 16S. Leur pouvoir antibiotique a été évalué par des expériences d'antagonisme contre des souches pathogènes humaines et les résultats de l'activité antibiotique des extraits issus des monocultures sont consécutivement présentés. Au cours d'un travail de fractionnement bioguidé de l'extrait d'une souche active, nous présentons le travail de purification et d'identification structurale d'une molécule active isolée. Le fractionnement bioguidé est une méthode de chimie des produits naturels qui recherche à identifier la ou les molécules responsables de l'activité biologique au sein d'un extrait en fractionnant l'extrait par chromatographie afin de séparer chacune des molécules de l'extrait. Enfin, les résultats des expériences de cocultures de souches d'actinobactéries sont décrits, où l'objectif est d'identifier des métabolites présents uniquement dans les cocultures en utilisant les réseaux moléculaires générés à partir des données spectrales des extraits de cocultures.

II. Partie Expérimentale

A. Collectes des fourmis

Toutes les fourmis ont été collectées lors de plusieurs missions en 2014 et 2015 en Guyane : sur les sentiers autour de Cayenne (Loyola, Rorota), autour de Kourou (jardin de Frédéric), à Cacao (sentier Molokoi), à Montsinery (sentier Crique des 3 rois), à Apatou (saut Herminia), à St Georges (saut Maripa), à la montagne de Kaw et autour du barrage de Petit Saut. Les espèces étudiées au sein de cet échantillonnage colonisent différents types d'habitat terricole, arboricole et myrmécophyte.

Le prélèvement des fourmis est réalisé à l'aide de pinces stérilisées en les chauffant au rouge au bec benzène puis trempées dans de l'alcool à 70% entre chaque récolte d'individus. Les fourmis sont ensuite introduites dans des tubes stériles directement sur le terrain. Une dizaine d'individus appartenant à la caste des ouvrières sont collectées par colonie. Six individus sont utilisés ensemble pour le lavage de la cuticule, et deux à trois individus sont conservés pour l'identification des espèces à posteriori.



Figure 19 : 1) entrée du nid de *Paraponera clavata* (auteur : Sacha Cataldo), 2) prélèvement sur le terrain (auteur : Christophe Duplais) et 3) mise en culture sur boîtes avec le laboratoire mobile (auteur : Christophe Duplais).

B. Isolement des actinobactéries

1) *Lavage*

Les fourmis sont ramenées le plus rapidement possible au laboratoire. Si le lieu éloigné de la collecte ne le permet pas, un laboratoire de microbiologie de terrain (becs électriques pour travailler dans une zone de stérilité) est mis en place pour traiter les échantillons sur le lieu de collecte. Afin de faciliter leur manipulation, les fourmis sont endormies à -20°C quelques minutes, ce qui permet d'éviter le relargage de venins ou d'autres substances issues des glandes

exocrines sous l'effet du stress pendant le lavage cuticulaire. Six fourmis de la même colonie sont insérées dans un tube contenant 1 mL d'une solution tampon (stérilisée par filtration à 0.20 µm) à l'aide de pince stérilisées. La solution tampon comprend différents sels minéraux qui évitent la détérioration des bactéries par choc osmotique et un surfactant (tween 80) en faible concentration ce qui optimise le décollement des micro-organismes de la cuticule (solution tampon : 0,7g de K₂HPO₄, 0,5 g de MgSO₄, 0,3g de KH₂PO₄, 0,01g de FeSO₄, 0,001g de ZnSO₄, 0,05% de tween 80, eau milli Q 1 L) (Sen *et al.* 2009). Les fourmis dans la solution sont ensuite mélangées 30 secondes afin d'optimiser le décollement mécanique des bactéries présentes sur la cuticule.

2) *Mise en culture*

100 µL de la suspension bactérienne issue du lavage cuticulaire sont étalés sur une boîte de Pétri. Trois types de milieux gélosés sont ensemencés, le milieu Chitine, Rorota et Actinomycetes, et chacun de ces trois ensemencements est répliqué une fois. Les boîtes sont incubées à température ambiante (28°C-30°C) pendant plusieurs semaines, et observées régulièrement afin de suivre l'apparition des actinobactéries.

Le milieu Chitine : pauvre en nutriments (pas de sucres ou de protéines) contenant des sels minéraux et de la chitine. La chitine est le composant principal de la cuticule des fourmis, et des insectes en général. Certaines actinobactéries ont la propriété de dégrader la chitine en hydrolysant les sucres, en formant des cercles clairs autour des colonies (Marsh *et al.* 2013). C'est une technique d'identification des actinobactéries. Ce milieu favorise également le développement des actinobactéries qui ont une croissance lente par rapport aux autres micro-organismes (Hanshew *et al.* 2015) (milieu Chitine : 5,33g de chitine, 0,76g de K₂HPO₄, 0,36g de KH₂PO₄, 0,24g de MgSO₄, 0,01g de FeSO₄, 0,001g de ZnSO₄, pH= 8,5 NaOH, eau milli Q 1L).

Le milieu Rorota : inspiré du milieu chitine et mis en place au laboratoire, il est composé de chitine et de minéraux issus d'un extrait de terre prélevée en forêt pour imiter la composition des minéraux du sol où fourragent les fourmis en Guyane (Zucchi *et al.* 2011) (milieu Rorota : 200g de terre, à ébullition 5 min, filtration coton. 100 mL filtrat, 5,33g de chitine, pH= 8,5 NaOH, eau milli Q 1 L).

Le milieu Actinomycetes : spécifique au développement des actinobactéries. Actinomycetes Broth, Vegitone, fabricant : Fluka Analytical (Lot BCBG 3312V).

3) Isolement des actinobactéries :

A l'aide de critères morphologiques principalement, certaines souches sont choisies et isolées sur un milieu gélosé riche favorisant la croissance des actinobactéries, le milieu YMEA (Marsh *et al.* 2013) (milieu YMEA : 4g de yeast extract, 10g de malt extract, 10g de dextrose, eau milli Q 1 L). Les actinobactéries sont reconnaissables à leur aspect duveteux qui apparaît avec le temps, elles sont incluses dans la gélose, elles sentent « la terre » (production de géosmine) (Figure 20) et créent des cercles d'inhibition clairs autour d'elles en dégradant la chitine (Figure 21).

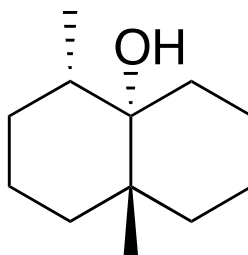


Figure 20 : Structure chimique de la géosmine.

Les souches isolées sont observées au microscope afin d'éliminer des levures ou champignons isolés par erreur. Les actinobactéries ressemblent fortement à des moisissures et il est parfois difficile de les différencier avec une simple observation macroscopique. Toutes les cultures sont incubées à 30°C, et des photos sont prises pour chacune. Les souches d'actinobactéries sont conservées au congélateur à -80°C, avec deux aliquotes par souche constituées de plusieurs öses d'actinobactéries mises dans une solution de conservation à base de glycérol (11,1 g de CaCl₂, 150 mL de glycérol, 850 mL d'eau) ce qui permet à la souche d'être conservée durablement.

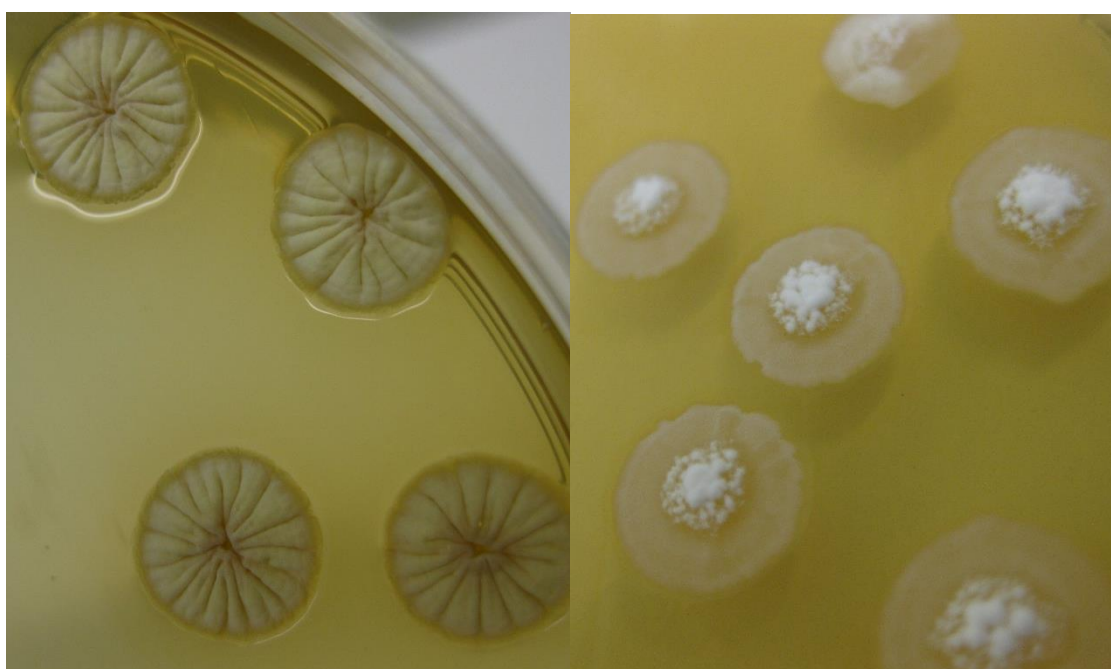
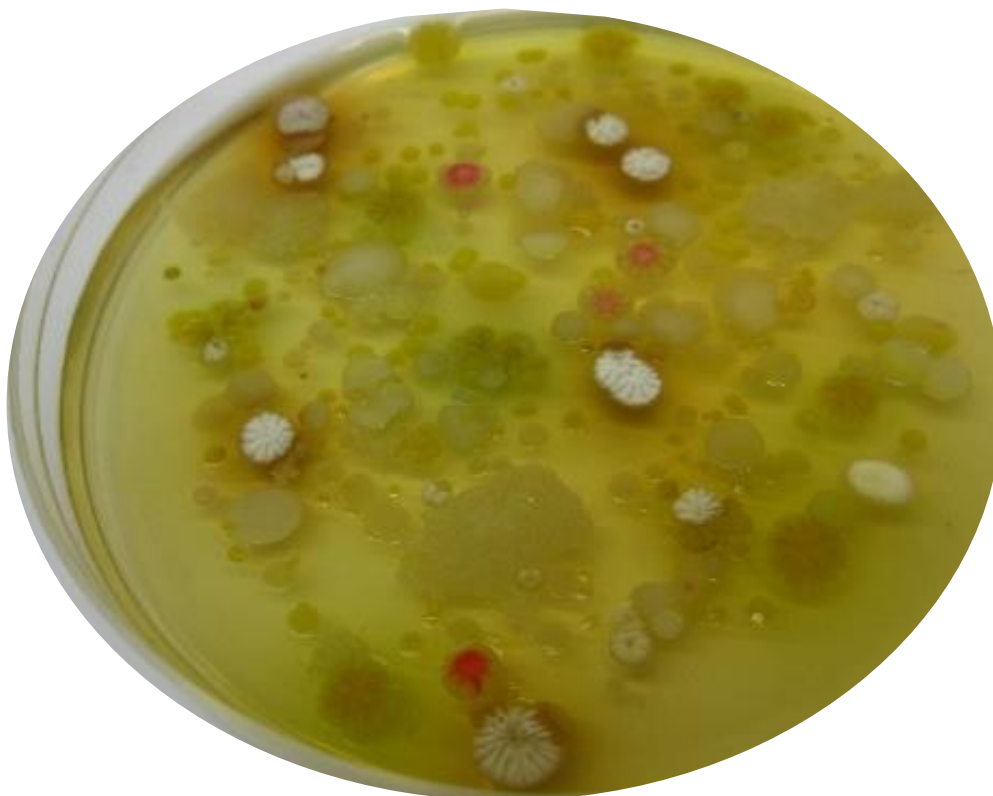


Figure 21 : Photos de boîtes de Pétri avec 1) plusieurs souches bactériennes, 2) la souche isolée CBCu566 et 3) CBCu788.

C. Identification par séquençage de l'ARNr 16S

L'ADN codant pour l'ARNr 16S des souches identifiées comme actinobactérie est amplifié puis séquencé par la méthode Sanger. L'ADN des souches sélectionnées est extrait à l'aide d'un automate MagMAX et du kit MagMAX-96 DNA Multi-Sample Kit où toutes les étapes ont été respectées, cette technologie permet de réaliser 96 extractions simultanément en plaque et est économe en temps par rapport aux techniques manuelles. Les acides nucléiques accrochés à des billes magnétiques (DNA Binding Bead Mix) sont déplacés par les aimants de l'automate d'étapes en étapes : lyse bactérienne (protéinase K), lavages des acides nucléiques et élution des ADN.

Les amorces universelles utilisées pour l'amplification par PCR sont le couple 27F (AGA GTT TGA TCM TGG CTC AG) et 1492R (CGG TTA CCT TGT TAC GAC TT) (Frank *et al.* 2008) utilisés en première intention pour l'identification des souches bactériennes et qui amplifient un fragment d'environ 800 pb de long. L'amplification PCR est effectuée dans un volume final de 25 µL. Le mélange PCR est composé de 1,25 µL d'ADN, de 1,25 µL (0,5 µM) de chaque amorce, de 0,37 µL (150 µM) d'un mélange équivalent contenant les dNTP, 1,5 µL (1,5 Mm) de MgCl₂, 0,25 µL d'AmpliTaq Gold Polymerase (Applied Biosystems, Foster City, CA, USA), 2,5 µL du tampon AmpliTaq Gold Buffer et 16.63 µL d'eau ultra pure. L'amplification par PCR consiste en une étape de dénaturation initiale à 94°C pendant 10 min suivi de 30 cycles à 94°C pendant 30s, 55°C pendant 30 s et 72°C pendant 30 s, et une étape d'extension finale à 72°C pendant 10 min. Une vérification des produits PCR est réalisée avec la migration de 5 µL de produit PCR mélangés à 1,3 µL de tampon de charge 10X sur un gel d'agarose à 2%. 20 µL de produits PCR sont envoyés à la société Beckman Coulter Genomics au Royaume-Uni (<http://www.beckmangenomics.com/>) pour un séquençage SANGER en sens et anti sens. Les séquences sens et anti-sens sont assemblés puis nettoyées avec le logiciel Mega 6 (Tamura *et al.* 2013). Des séquences d'environ 1200 pdb sont obtenues et utilisées pour l'identification des souches isolées à l'aide du serveur EzTaxon qui contient une base de référence de séquence du gène de l'ARN 16S (Kim *et al.* 2012) (<http://www.ezbiocloud.net/eztaxon>).

D. Pouvoir antibiotique des souches d'actinobactéries isolées :

L'objectif ici est d'évaluer l'activité antagoniste des 43 souches isolées sur la cuticule de fourmi contre les bactéries pathogènes pour l'humain : *Escherichia coli*, *Enterococcus faecalis*, *Staphylococcus aureus* et *Pseudomonas aeruginos* (Table 3). Ces souches sont les souches classiquement testées lors de la prospection de molécules antibiotiques. Leurs différences de Gram et de leur caractéristique respiratoire permettent la recherche d'une activité

antibiotique à partir de bactéries représentatives de la diversité des Grams et des systèmes respiratoires existant chez les bactéries. La méthode cross-streak utilisée est un test de cocultures sur milieu solide permettant de mimer les interactions entre les micro-organismes isolées et les souches pathogènes sélectionnées. Afin de mieux caractériser le spectre d'action des actinobactéries isolées des cuticules de fourmi, des tests d'antagonismes contre une levure *Candida albicans* et un champignon entomopathogène *Beauveria bassiana* sont aussi réalisés.

1) Méthode cross-streak

La méthode cross-streak sur gélose consiste à inoculer perpendiculairement à la bande d'actinobactérie des bandes de « souche test » dont on observera la croissance par rapport à celle d'un témoin négatif (sans l'actinobactérie). Les « souches tests » sont les souches pathogènes pour l'homme ou la fourmi à partir desquelles l'activité d'inhibition des actinobactéries est observée (Table 3). L'aspect duveteux des actinobactéries à maturité ne permet pas d'obtenir des démarcations bien nettes lors de l'inoculation de la souche. Le test cross-streak a été modifié pour les besoins du screening d'actinobactéries bioactives. Ainsi l'inoculation de la souche testée est décalée de 1 cm par rapport au bord de la bande centrale de la souche d'intérêt. Afin d'éliminer les biais de lecture des zones d'inhibition un support modèle en papier millimétré est utilisé (Velho-Pereira & Kamat 2011). Le protocole détaillé est en Annexe 1.

Table 3 : Description des souches pathogènes choisies.

	Bactérie Gram négatif		Bactérie Gram positif		Champignon	Levure
Nom	<i>Escherichia coli</i> NCTC 13167	<i>Pseudomonas aeruginosa</i> ATCC 9027	<i>Staphylococcus aureus</i> ATCC 9144	<i>Enterococcus faecalis</i> CCM 25415	<i>Beauveria bassiana</i> DSM 1344	<i>Candida albicans</i>
Habitat	tractus digestif	ubiquitaire	la peau et les muqueuses	flore intestinale	le sol, les insectes	les muqueuses, tube digestif
Aérobic/anaérobic	aérobic fermentaire	aérobic oxydative	aérobic oxydative	anaérobic facultatif fermentaire	aérobic stricte	aérobic
Pathogénicité	gastro-entérites infection urinaires, méningites, ou sepsis	maladies nosocomiales, infection de l'œil, des plaies, des urines, gastro-intestinales et des poumons	maladies nosocomiales, intoxications alimentaires, infections localisées suppurées et cas extrêmes de septicémies	maladies nosocomiales, inflammations chroniques de l'intestin, endocardites, infections de la vessie	parasite un insecte en l'attaquant	candidose, érythème fessier, vaginite
Milieu de culture et température	LB, 30 °C	LB, 30 °C	LB, 30 °C	LB, 30 °C	PDA, 25 °C	Sabouraud, 30 °C

La lecture des résultats est faite à 48h à l'aide d'un papier calque millimétré (dessiné à partir du modèle ayant servi à la mise en culture des souches tests) et le nombre de carré de 1 mm de côté remplissant la zone d'inhibition est compté. Une bande de souche test non inhibée, par exemple celles présentes sur le témoin négatif, recouvre une surface de 150 mm². Le

pourcentage d'inhibition d'une actinobactérie pour chaque souche testée est obtenu avec cette formule : $\text{pourcentage d'inhibition} = (\text{surface zone inhibée}) / (\text{surface zone totale}) * 100$.

2) *Méthode broth dilution*

Les actinobactéries sont mise en culture liquide dans du milieu YMEA pendant 5 jours. Lors de cette fermentation, les actinobactéries se multiplient en produisant des métabolites libérés dans le milieu. L'extraction par un mélange de solvant organique du milieu permet d'obtenir un extrait comprenant les métabolites.

(a) Préparation des extraits

Le protocole d'extraction des métabolites en culture liquide s'inspire d'une méthode d'extraction des métabolites de champignons cultivés sur gélose solide (Bertrand et al. 2014). Les principales étapes sont 1) la fermentation de la souche en milieu liquide pendant 5 jours 2) l'arrêt de la prolifération bactérienne 3) la lyophilisation 4) l'extraction aux ultra-sons 5) l'élimination des sucres par lavage à l'eau de l'extrait sur SPE (Figure 22). Le protocole détaillé est en Annexe 2.

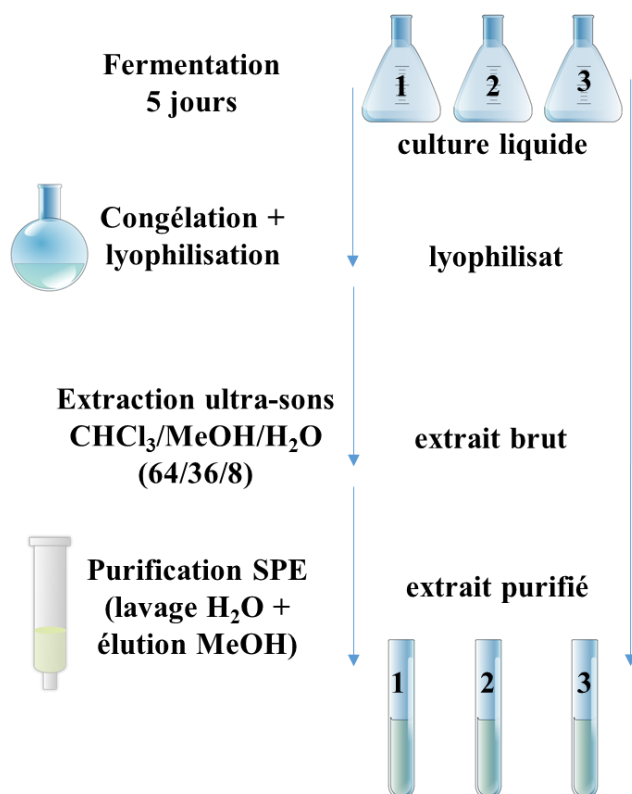


Figure 22 : Schéma bilan de la méthode d'extraction des métabolites.

(b) Pourcentage d'inhibition des extraits

Chaque extrait purifié est testé de manière identique dans 3 puits à 400 µg/mL contre quatre souches bactériennes pathogènes pour l'homme : *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus* et *Enterococcus faecalis*. Les tests antibactériens sont réalisés dans des plaques 96 puits suivant un plan de plaque optimisé (Wiegand *et al.* 2008). La lecture des résultats est réalisée par une mesure de l'absorbance à 620 nm des puits, ainsi le puits de contrôle avec la culture bactérienne se développant sans contrainte a un aspect trouble et une mesure d'absorbance plus élevée que les puits contenant un extrait inhibant la prolifération bactérienne présentant un aspect moins trouble. Les résultats de la lecture des plaques permettent de calculer le pourcentage d'inhibition de l'extrait à 400 µg/mL face aux bactéries pathogènes. Le protocole détaillé de ces tests antibactériens en culture liquide est en Annexe 3.

$$\text{Pourcentage d'inhibition de la croissance} = \left[1 - \frac{\text{DO}_{\text{drug}} - \text{DO}_{\text{milieu}}}{\text{DO}_{\text{parasite seul}} - \text{DO}_{\text{milieu}}} \right] \times 100$$

Figure 23 : Formule de calcul du pourcentage d'inhibition de la croissance bactérienne, DO drug = mesure d'absorbance du puits contenant la suspension bactérienne et l'extrait testé, DO milieu = mesure d'absorbance du milieu de culture stérile, DO parasite seul = mesure d'absorbance de la suspension bactérienne seule.

E. Isolement et identification d'une molécule chimique

1) Souche 497 : culture liquide, extraction, isolement

La souche 497 est sélectionnée pour effectuer un fractionnement bioguidé à partir d'une plus grande quantité d'extrait car son extrait inhibe à plus de 50% la croissance d'*Enterococcus faecalis* et de *Pseudomonas aeruginosa*.

Obtention de l'extrait : 20 mL de la pré-culture (souche pure introduite dans un erlenmeyer de 250 mL contenant 50 mL de milieu YMEA, agitation 48 heures à 30°C) sont introduits dans un erlenmeyer de 5 L contenant 1,5 L de milieu YMEA liquide stérile. 5 réplicats sont réalisés de manière identique. Les cultures sont agitées pendant cinq jours à 30°C. La prolifération bactérienne est arrêtée par le froid (congélation azote liquide) et les 7,5 L de culture liquide sont lyophilisés ce qui donne 120 g de lyophilisat. 12 g de lyophilisat sont extraits aux ultra-sons pendant 20 min dans 400 mL d'un mélange de solvant dichlorométhane/méthanol/eau (64/36/8). Cette opération est répétée 9 fois afin d'extraire tout le lyophilisat. 46 g de l'extrait

brut sont obtenus. Afin d'éliminer une partie du milieu de culture contenu dans l'extrait brut, les 46 g sont dissous dans 200 mL d'eau pure et lavés quatre fois avec 200 mL d'acétate d'éthyle. La phase organique est évaporée à l'évaporateur rotatif et résulte en un extrait purifié de 1,2 g.

Fractionnement bioguidé : une petite partie (200 mg) de l'extrait est fractionnée sur une colonne ouverte de silice C18 en utilisant un gradient acétonitrile (ACN)/(H₂O) (25:75 ; 50:50 ; 75:25 ; 100:0) terminé d'une élution finale à l'acétate d'éthyle. L'extrait est déposé sec imprégné sur de la silice C18 en tête de colonne. 16 fractions sont séchées et leurs CI₅₀ (concentration inhibitrice de 50%) sont déterminées avec une gamme de concentration de 400 µg/mL à 3,1 µg/mL. Les fractions F4 à F6 présentant des CI₅₀ intéressantes sont analysées en HPLC (Varian 920-LC) couplé à un détecteur UV (détecteur à barrette de diode) à 210 nm. La colonne est une C18 15×4,6 ; 5 µm Discovery. Les pics majoritaires de ces fractions sont par analogie des temps de rétention repérés sur le chromatogramme de l'extrait non fractionné. Une méthode HPLC semi-préparative (Waters 600 pump, 2487 dual absorbance, 600 controller) est adaptée et 1 g d'extrait est injecté en 10 fois pour obtenir 6,5 mg d'une molécule appelée **1**.

2) Détermination structurale : RMN, Spectrométrie de masse et Dichroïsme circulaire

Les expériences de RMN ont été réalisées sur un spectromètre Varian (400 MHz).

Analyses de spectroscopie RMN : Les spectres RMN ont été réalisés sur un spectromètre Varian (400 MHz) et l'échantillon est solubilisé dans du diméthylsulfoxyde deutéré (DMSO-*d*₆). Ci-dessous la liste des séquences utilisées et les paramètres des expériences associées.

- ¹H : mesure directe de la relaxation du spin du noyau hydrogène. 64 scans ; nombre d'incréments = 512 ; temps de relaxation = 10 s.

- gCOSY (*gradient-selected COrrrelation SpectroscopY*) : met en évidence les couplages entre protons vicinaux et adjacents (²*J*, ³*J*) représentés par des tâches de corrélation et permet de connaître l'enchaînement des protons voisins. Cette séquence utilise des gradients d'impulsion pour éliminer les magnétisations indésirables et ainsi obtenir des spectres avec une excellente résolution en seulement quelques scans. 4 scans ; nombre d'incréments = 512 ; temps de relaxation = 3 s.

- gHSQCAD (*Adiabatic gradient-selected Heteronuclear Single Quantum Correlation*) : mesure le couplage entre un proton et le carbone auquel il est relié en utilisant des gradients d'impulsions à 180° du ¹³C. 8 scans ; nombre d'incréments = 512 ; temps de relaxation = 2 s.

- gHMBC (*gradient-selected Heteronuclear Multiple Bond Correlation*): permet d'établir les corrélations proton-carbone à longue distance. 8 scans ; nombre d'incréments = 512 ; temps de relaxation = 2 s.

Analyses de spectrométrie de masse : Le composé **1** isolé est analysé sur un système Agilent LC-MS (Agilent 1260 infinity HPLC, Agilent 6530 Q-TOF-MS, ESI, mode positif) avec une colonne Agilent extend C18 (150 × 2,1 mm; i.d. 3,5 µm), avec un débit à 0,9 mL/min et un gradient linéaire 98/2 A/B à 0/100 A/B en 7 min (A: H₂O à 0,1% d'acide formique, B: acétonitrile à 0,1% d'acide formique). Les conditions de la source ESI sont : température du capillaire=320°C, voltage de la source=3,5 Kv, gaz=12 L/min. En mode positif, purine C₅H₄N₄ [M+H]⁺ ion (m/z 121,050873) et hexakis (1H, 1H, 3H-tetrafluoropropoxy)-phosphazine C₁₈H₁₈F₂₄N₃O₆P₃ [M+H]⁺ ion (m/z 922,009798) sont utilisés comme des masses de contrôle interne.

Analyses de dichroïsme circulaire : Le dichroïsme circulaire repose sur la capacité d'une molécule possédant un ou plusieurs centres chiraux (optiquement actif) d'absorber sélectivement la lumière circulaire polarisée. La polarisation spécifique peut être circulaire droite ou circulaire gauche. Le spectre dichroïque correspond à la différence d'absorbance entre ces deux types de lumière pour chaque longueur d'onde. Les analyses expérimentales effectuées par spectromètre à dichroïsme circulaire ainsi que la modélisation théorique ont été réalisées par Grégory Genta-Jouve à l'Université Paris Descartes.

F. Les cocultures

Les étapes de préparation des extraits et les tests antibiotiques sont réalisés de manière similaire aux protocoles détaillés en Annexe 2 et Annexe 3. Les quelques ajustements spécifiques à la réalisation des cocultures lors de la fermentation de cinq jours sont également détaillés en Annexe 2.

G. Analyse de spectrométrie de masse : réseau moléculaire

Analyses de spectrométrie de masse : Les extraits de monoculture et coculture sont analysés sur un système Agilent LC-MS (Agilent 1260 infinity HPLC, Agilent 6530 Q-TOF-MS, ESI, mode positif) avec une colonne Agilent extend C18 (150 × 2,1 mm; i.d. 3,5 µm), avec un débit à 0,8 mL/min et un gradient linéaire 98/2 A/B à 0/100 A/B en 7 min (A: eau à 0,1% d'acide formique, B: acétonitrile à 0,1% d'acide formique). Les conditions de la source ESI sont : température du capillaire=320°C ; voltage de la source=3,5 Kv ; gaz=12 L/min. En mode positif, les ions sont sélectionnés dans la fenêtre m/z 100-1700, puis trois expériences indépendantes à trois énergies de collision données (20 ; 40 ; 80 eV) sont effectuées pour

chacun des trois pics majoritaires d'un scan. En mode positif, purine C₅H₄N₄ [M+H]⁺ ion (m/z 121,050873) et hexakis (1H, 1H, 3H-tetrafluoropropoxy)-phosphazine C₁₈H₁₈F₂₄N₃O₆P₃ [M+H]⁺ ion (m/z 922,009798) sont utilisés comme des masses de contrôle interne.

Construction des réseaux : Les réseaux moléculaires sont générés à partir de données HPLC-HRMS/MS des extraits d'actinobactéries en monoculture et en coculture. Les données sont converties à partir du format standard .d (Agilent) au format .mgf avec le logiciel MassHunter Workstation (Agilent). Les réseaux moléculaires sont réalisés sur la plateforme internet GNPS (<http://gnps.ucsd.edu>) (Wang *et al.* 2016). Les paramètres utilisés sont minimum pairs cos = 0,7 ; precursor ion mass tolerance = 2,0 Da ; fragment ion mass tolerance = 0,5 ; network topK = 10 ; minimum matched fragment ions = 6 ; minimum cluster size = 2. Les paramètres utilisés pour la comparaison aux banques spectrales de GNPS (bases de données) sont library search min matched peaks = 6 ; score threshold = 0,7. Ainsi, toutes les identifications affichées sur le réseau sont gardées si le score de cosinus est supérieur à 0,7 et si au moins 6 pics sont communs entre les données expérimentales et celles des banques spectrales. Les spectres MS² ayant une forte similarité spectrale (CS > 0,7) sont représentés sous forme de nœuds (*node*) et les liaisons entre ces nœuds (*edge*) apparaissent sur les réseaux moléculaires dès lors que le score CS est supérieur à 0,7 et qu'au moins six ions communs sont détectés. L'épaisseur des connections entre les nœuds est proportionnelle à la valeur du CS. Le rapport m/z de l'ion précurseur est inséré au centre du nœud et l'identification proposée, selon les critères indiqués ci-dessus, est inscrite à côté lorsqu'il y en a une. Les données des réseaux moléculaires générés sont visualisées avec le logiciel Cytoscape version 3.4 (<http://cytoscape.org/>).

III. Résultats et discussions

A. Souches d'actinobactéries isolées

Parmi les 43 souches d'actinobactéries isolées sur la cuticule de 15 espèces de fourmis, on dénombre 31 *Streptomyces*, 5 *Kitasatospora*, 4 *Tsukamurella*, 1 *Actinophytocola*, 1 *Leucobacter* et 1 *Arthrobacter* (Table 4). L'identification des 43 souches est obtenue par alignement et comparaison des séquences de gène 16S ARN ribosomal avec la base de données [Eztaxon](#) en accès libre (Chun *et al.* 2007; Kim *et al.* 2012). Le pourcentage de similarité et la référence (référéncée par son numéro d'accession) de la séquence la plus proche sont indiqués dans la Table 4. Seules quatre souches ont des similarités inférieures à 97% par rapport à leur souche la plus proche dans la base de données par conséquent leur identification n'est pas validée (Tindall *et al.* 2010).

Table 4 : Liste des espèces de fourmis (habitat, lieux et dates de collectes) et identification des 43 souches d'actinobactéries isolées (souche la plus proche et pourcentage de similarité selon EzTaxon).

N° de la souche isolée	Espèce de fourmi	Mode de vie	Localisation et date de collecte	Souche la plus proche	N° de la souche	Pourcentage de similarité (%)
51-111	<i>Camponotus femoratus</i>	jardin de fourmi	Petit Saut, 04/15	<i>Streptomyces griseoluteus</i>	AB184363	98.33
788	<i>Cephalotes atratus</i>	arboricole	Kourou, 03/14	<i>Streptomyces purpurascens</i>	AB184859	99.05
32-51				<i>Streptomyces chattanoogensis</i>	AJ621611	98.79
32-41	<i>Dolichoderus atelabroides</i>	terricole	Cacao, 03/15	<i>Streptomyces gramineus</i>	HM748598	99.93
32-62				<i>Streptomyces yokosukanensis</i>	DQ026652	99.13
48-31				<i>Kitasatospora cheerisanensis</i>	KK853997	99.49
48-212				<i>Kitasatospora gansuensis</i>	AY442265	99.27
48-43	<i>Eciton burchelli</i>	terricole	Sentier trois Rois, 04/15	<i>Kitasatospora saccharophila</i>	AB278568	93.19
48-14				<i>Streptomyces acidiscabies</i>	D63865	99.64
48-13				<i>Streptomyces pactum</i>	AB184398	99.63
902			Apatou, 04/14	<i>Streptomyces xanthocidicus</i>	AB184427	99.28
46-41			Sentier trois Rois, 04/15	<i>Kitasatospora phosalacinea</i>	JNWZ01000148	99.63
46-51				<i>Streptomyces purpurascens</i>	AB184859	99.13
555				<i>Streptomyces coeruleus</i>	AB184122	99.79
5510	<i>Ectatomma tuberculatum</i>	terricole		<i>Streptomyces hiroshimensis</i>	AB184802	96.74
564			Cayenne, 01/14	<i>Streptomyces luteosporus</i>	AB184607	98.92
581				<i>Streptomyces prunicolor</i>	BARF01000082	97.50
5513				<i>Streptomyces thioluteus</i>	AJ781360	99.28
566				<i>Tsukamurella pulmonis</i>	X92981	99.64
444	<i>Gnamptogenys sp</i>	terricole	Cayenne, 01/14	<i>Leucobacter humi</i>	KC818288	99.78
447				<i>Tsukamurella tyrosinosolvens</i>	Y12246	99.71
4916				<i>Streptomyces coacervatus</i>	AB500703	98.85
492	<i>Labidus sp.</i>	terricole	Cayenne, 01/14	<i>Streptomyces hiroshimensis</i>	AB184802	93.47
497				<i>Streptomyces thioluteus</i>	AJ781360	99.28
518			Cayenne, 01/14	<i>Arthrobacter protophormiae</i>	X80745	98.28
7821	<i>Neoponera commutata</i>	terricole		<i>Streptomyces avellaneus</i>	AB184413	99.78
7819			Kourou, 03/14	<i>Streptomyces lunalinharesii</i>	DQ094838	99.28
7811				<i>Streptomyces pulveraceus</i>	AJ781377	99.71
38-11	<i>Odontomachus biumbonatus</i>	terricole	Cacao, 03/15	<i>Streptomyces panaciradicis</i>	KF971876	98.24
38-22				<i>Streptomyces xanthocidicus</i>	AB184427	98.44
9318	<i>Odontomachus haematodus</i>	terricole	Kaw, 04/14	<i>Streptomyces gramineus</i>	HM748598	99.85
40-62				<i>Kitasatospora gansuensis</i>	AY442265	98.62
40-11			Cacao, 03/15	<i>Streptomyces hokutonensis</i>	AB808756	99.49
43-62	<i>Odontomachus hastatus</i>	terricole		<i>Streptomyces kunmingensis</i>	AB184597	98.92
53-12			Petit Saut, 04/15	<i>Actinophytocola corallina</i>	AB511316	99.85
53-41				<i>Streptomyces kunmingensis</i>	AB184597	98.76
44-313	<i>Paraponera clavata</i>	terricole	Kaw, 03/15	<i>Streptomyces xanthocidicus</i>	AB184427	98.25
41-52	<i>Pheidole minutula</i>	myrmécophyte	Cacao, 03/15	<i>Streptomyces anandii</i>	AY999803	98.84
62-41			Petit Saut, 04/15	<i>Streptomyces griseoluteus</i>	AB184363	98.69
417	<i>Pseudomyrmex penetrator</i>	myrmécophyte	St Georges, 04/14	<i>Streptomyces michiganensis</i>	AB184153	99.93
45-21			Kaw, 03/15	<i>Tsukamurella inchonensis</i>	X85955	93.76
59-11	<i>Solenopsis sp.</i>	terricole	Petit Saut, 04/15	<i>Streptomyces griseofuscus</i>	AB184206	99.93
34-61			Cacao, 03/15	<i>Tsukamurella tyrosinosolvens</i>	Y12246	99.78

B. Pouvoir antibiotique des actinobactéries isolées :

Les résultats de tests d'inhibition présentés dans la Table 5 sont classés en quatre catégories pour faciliter l'interprétation des résultats : inhibition totale (100% d'inhibition), inhibition forte ($\geq 50\%$), inhibition faible (entre 15% et 50%) et absence d'inhibition ($< 15\%$).

1) Antagonismes aux quatre bactéries pathogènes :

Les résultats du test cross-streak montrent que 22 souches d'actinobactéries sur 43 (51%) présentent une activité antagoniste à au moins une des souches pathogènes testées. Plus

précisément, 20 souches inhibent la croissance de bactéries *Staphylococcus aureus* et 14 souches inhibent celle de *Enterococcus faecalis*. Ces deux espèces sont Gram positives. D'autre part 8 souches inhibent *Escherichia coli* et 7 souches inhibent *Pseudomonas aeruginosa*, toutes deux Gram négatives. Notons que six actinobactéries présentent une inhibition totale pour *Staphylococcus aureus*, dont trois inhibent également totalement *Enterococcus faecalis*. Une seule souche isolée inhibe totalement la croissance de *Pseudomonas aeruginosa*, et la plupart des inhibitions des bactéries Gram négatif sont de la catégorie d'inhibition faible.

Indépendamment des valeurs d'activités d'antagonisme observées, les souches d'actinobactéries inhibent majoritairement les souches pathogènes Gram positives. Ce résultat a déjà été observé dans la littérature (Mearns-Spragg *et al.* 1998) et rappelons que les actinobactéries sont elles aussi des bactéries de type Gram positif. Les auteurs suggèrent l'existence d'un quorum sensing amplifié entre les bactéries Gram positif. Le quorum sensing correspond au mécanisme de synchronisation de l'expression (ou de la répression) de gènes particuliers, au sein d'une population ou d'une communauté bactérienne et est fonction de la densité de cette population/communauté. Le quorum sensing implique la capacité des bactéries à communiquer avec leurs voisins via des signaux moléculaires. Ainsi les actinobactéries reçoivent et émettent plus de métabolites de communication, ou antibiotiques lorsque qu'elles sont face à des bactéries de type Gram positif.

2) Antagonismes à *Candida albicans* et *Beauveria bassiana* :

Deux autres souches pathogènes d'origine fongique ont été choisies au cours de cette étude afin de mieux évaluer le spectre d'action des 43 actinobactéries isolées : *Candida albicans*, une levure pathogène pour l'Homme et *Beauveria bassiana*, champignon entomopathogène très répandu chez les insectes et notamment chez les fourmis (Reber & Chapuisat 2011).

D'un point de vue général, les résultats de la table 4 montrent que 21 souches d'actinobactéries ont des activités antifongiques. 14 de ces souches ont des activités antagonistes fortes et inhibent totalement la croissance de l'une ou l'autre des souches fongiques ou des deux souches fongiques. Notons que 4 souches d'actinobactéries ont une activité antifongique forte alors qu'elles ne présentent pas d'activité bactérienne. L'observation inverse n'est pas constatée.

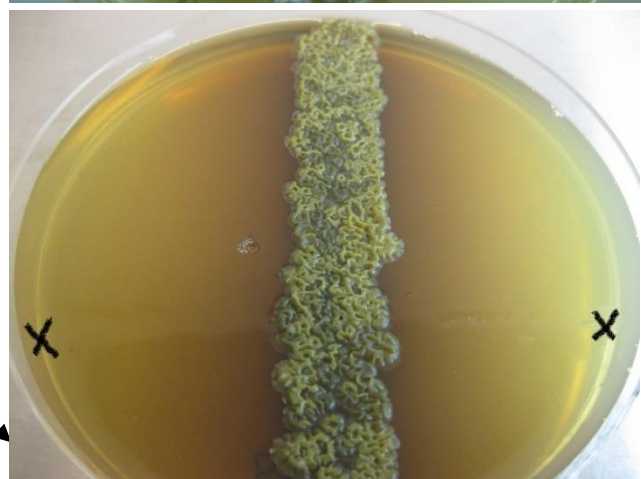
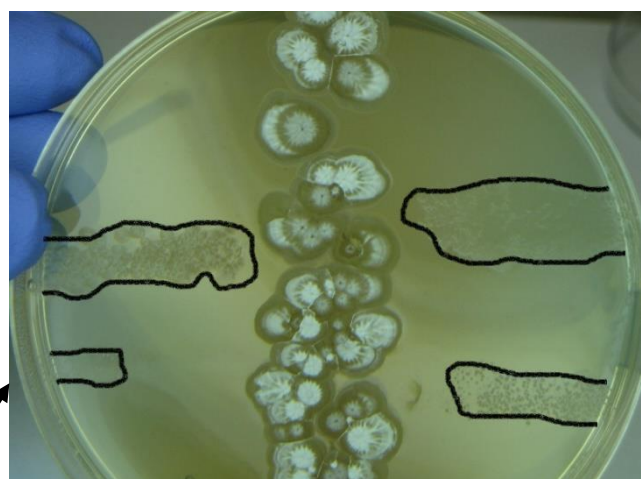
39% (17/43) des actinobactéries inhibent *Beauveria bassiana*. Ce résultat est intéressant et renvoi au rôle supposé de défense des actinobactéries sur la cuticule des fourmis. Ainsi la présence d'actinobactéries antagonistes aux bactéries pathogènes mais surtout à une souche

fongique entomopathogène renforce l'hypothèse du rôle défensif de ces actinobactéries pour la fourmi. D'un point de vue pharmacologique, les souches présentant des antagonismes multiples, c'est-à-dire aux deux types de Gram et aux souches fongiques, sont intéressantes pour la prospection de métabolites antibactériens à spectre large. Néanmoins, la prospection de métabolites antibactériens à spectre réduit est également d'un grand intérêt car leur mécanisme d'action peut être plus spécialisé et permettrait de limiter l'apparition de résistance.

L'évaluation de l'activité antimicrobienne reposant sur des tests d'antagonismes est couramment utilisée dans les études d'activité biologique par les biologistes, cependant cette approche est largement moins utilisée par les chimistes qui à la place évaluent l'activité biologique des extraits directement. Bien que les tests d'activité biologique des extraits réduisent le champ des possibles, surtout lors de la production de métabolites impliqués dans la compétition, il est nécessaire pour les chimistes d'avoir à disposition un extrait dont ils sont sûr qu'il contient le ou les métabolites actifs afin de les isoler pour les identifier. Il est difficile de savoir si le test d'activité d'antagonisme sur gélose permet d'identifier plus les souches actives que les tests d'activité biologique des extraits. Les mécanismes d'inhibition de croissance impliqués sont différents et conduisent à des résultats contrastés dans les différents tests. Dans une étude sur des souches isolées des milieux marins écossais (Burgess *et al.* 1999), les auteurs assurent avoir détectés plus de souches actives que la proportion admise lors de l'isolation de souches actives avec les tests d'activité biologique des extraits, notamment car la production des métabolites antibiotiques est stimulée par l'interaction sur le milieu solide de culture via la présence des souches bactériennes pathogènes humaines. Néanmoins l'identification des métabolites ou peptides associés à ces activités reste toujours délicate dans tous les cas.

Table 5 : Pourcentage d'inhibition des 43 souches d'actinobactéries contre 6 souches pathogènes, inhibition totale = +++ (rouge), inhibition forte entre 50% et 100% = + (orange) et inhibition faible entre 15% et 50% = +/- (jaune) et absence d'inhibition = - (blanc)

Code souche	<i>Staphylococcus aureus</i>	<i>Enterococcus faecalis</i>	<i>Pseudomonas aeruginosa</i>	<i>Escherichia coli</i>	<i>Beauveria bassiana</i>	<i>Candida albicans</i>
32-41	-	-	-	-	-	-
32-51	+	-	-	+/-	+++	+++
32-62	-	-	-	-	+++	+++
34-61	-	-	-	-	-	-
38-11	+/-	-	-	-	-	-
38-22	+	+	+/-	+	-	-
40-11	-	-	-	-	+	-
40-62	-	-	-	-	-	-
41-52	-	-	-	+++	+/-	-
43-62	+++	+++	+++	-	+/-	-
44-313	+	+/-	-	-	+++	-
45-21	-	-	-	-	-	-
46-41	-	-	-	-	-	-
46-51	+/-	-	-	+++	+++	-
48-13	-	-	-	-	-	-
48-14	-	+/-	-	+/-	-	-
48-212	-	-	-	-	-	-
48-31	-	-	-	-	-	-
48-43	-	-	-	+++	+	-
51-111	-	-	-	-	-	-
53-12	-	-	-	-	-	-
53-41	-	-	-	-	-	-
59-11	-	-	-	+++	+++	-
62-41	+	-	-	-	-	-
444	-	-	-	-	-	-
447	+/-	-	+/-	-	-	-
7811	+	-	-	+/-	+/-	-
7819	+	+	-	+/-	+++	+++
7821	+/-	-	+/-	+/-	+/-	+
518	-	+/-	-	+++	-	-
492	+++	+/-	+/-	+	+/-	+++
497	+	+/-	-	+/-	+++	+
4916	-	-	-	-	-	-
555	+	+/-	-	-	+	-
5510	+++	+/-	+/-	+/-	+	+++
5513	+++	+++	+/-	+	+++	+++
564	+	+	-	-	-	-
566	-	-	-	-	-	-
581	+/-	+/-	-	-	+/-	-
788	-	-	-	-	-	-
902	+++	+++	-	-	+++	+++
9318	-	-	-	-	-	+
417	+++	-	-	-	+/-	-



C. Activité antibiotique des métabolites extraits des souches d'actinobactéries

Dans un deuxième temps nous avons déterminé l'activité antibactérienne des extraits de culture des souches d'actinobactéries isolées. Chaque extrait purifié est testé de manière identique dans 3 puits à 400 µg/mL contre les bactéries : *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus* et *Enterococcus faecalis*. Les résultats des inhibitions sont présentés dans la (Table 6) ci-dessous et utilisent la même échelle de couleur que précédemment. Les résultats d'inhibitions sans l'échelle de couleur sont en (Annexe 4). Les résultats de tests d'inhibition sont classés en quatre catégories pour faciliter l'interprétation des résultats : inhibition totale (100% d'inhibition), inhibition forte ($\geq 50\%$), inhibition faible (entre 15% et 50%) et absence d'inhibition.

13 extraits provenant des 43 souches différentes d'actinobactéries isolées (30%) présentent une activité antibiotique et 4 extraits ont une inhibition totale de la croissance de *Pseudomonas aeruginosa*. Parmi ces derniers, 2 inhibent également à 100% la croissance d'*Enterococcus faecalis*. On remarque lors de l'évaluation des activités biologiques des extraits réalisés à partir de cultures d'actinobactéries qu'un nombre inférieur de souche montre des activités antibactériennes qu'avec les résultats obtenus lors du test d'antagonisme. Bien qu'il n'y ait pas de relation quantitative et qualitative entre les deux tests, nous constatons que les extraits testés inhibent moins fortement la prolifération bactérienne des 4 souches pathogènes humaines que les souches d'actinobactéries elles-mêmes. Une explication pourrait être que les extraits réalisés à partir de monoculture en milieu liquide sans compétiteur/pathogène ne contiendraient pas – ou en faible quantité- l'ensemble des molécules synthétisées lorsque la souche se développe en coculture. Ceci supporte donc l'hypothèse d'une activité antimicrobienne induite par la compétition. De plus, l'extraction du milieu de culture selon notre protocole et notamment l'étape où l'on lave à l'eau l'extrait pour éliminer les résidus du milieu de culture pourrait avoir également éliminée des molécules très hydrophiles produites par la souche. Cependant, pour la souche 497 où le fractionnement bioguidé a été effectué, la solution aqueuse issue du lavage de l'extrait brut a été testée et ne montrait aucune activité antibiotique, de plus l'analyse par HPLC UV de cette solution aqueuse présentait uniquement un pic très important au début du chromatogramme correspondant probablement à des sucres dont l'élution sur la colonne est rapide.

Il est intéressant de noter qu'aucun extrait n'inhibe la prolifération bactérienne de *Staphylococcus aureus* alors que c'est la souche la plus inhibée par les actinobactéries lors des tests d'antagonismes cross-streak. Peut-être que les concentrations de métabolites sont beaucoup trop faibles dans l'extrait pour avoir une activité, ou peut-être qu'un mécanisme de

stress a effectivement été mis en place sur le milieu solide et a conduit à la production de métabolites antibiotiques. L'autre souche pathogène de type gram positif *Enterococcus faecalis* est inhibée par 12 extraits. Précédemment cette souche était inhibée par 14 souches d'actinobactéries lors des tests d'antagonismes et 6 de ces souches présentent des extraits actifs. Dans ce cas les deux tests d'évaluation d'activité antibactérienne sont complémentaires dans une démarche pharmacologique, et l'évaluation de l'activité de l'extrait est indispensable si un fractionnement bioguidé est envisagé. En effet, au regard du temps nécessaire à la réalisation d'un fractionnement bioguidé il est important de s'assurer que les souches d'intérêt, et surtout leur extrait soit actif contre les pathogènes. Pour le chercheur en chimie des produits naturels, l'objectif principal est d'isoler et de caractériser le ou les métabolites actifs.

Table 6 : Pourcentage d'inhibition des extraits purifiés des 43 souches d'actinobactéries contre quatre souches tests pathogènes, inhibition totale = +++ (rouge), inhibition forte entre 50% et 100% = + (orange) et inhibition faible entre 15% et 50% = +/- (jaune) et absence d'inhibition = - (blanc). Photos d'actinobactérie après 5 semaines, de haut en bas : CBCu555, CBCu497 et CBCu782.

Code souche	Staphylococcus aureus	Enterococcus faecalis	Pseudomonas aeruginosa	Escherichia coli
32-41	-	-	-	+/-
32-51	-	-	-	-
32-62	-	-	-	-
34-61	-	-	-	-
38-11	-	-	-	-
38-22	-	-	-	-
40-11	-	-	+/-	-
40-62	-	-	-	-
41-52	-	+++	+++	-
43-62	-	+	+++	-
44-313	-	-	-	-
45-21	-	-	-	-
46-41	-	+/-	-	-
46-51	-	-	-	+/-
48-13	-	-	-	-
48-14	-	+/-	+/-	-
48-212	-	+/-	-	-
48-31	-	+/-	-	-
48-43	-	-	-	-
51-111	-	-	-	-
51-62	-	-	-	-
53-12	-	-	-	-
53-41	-	+/-	+/-	-
59-11	-	-	-	-
62-41	-	-	-	-
444	-	-	-	-
447	-	-	-	-
7811	-	-	-	-
7819	-	-	-	+/-
7821	-	-	-	-
518	-	-	-	-
492	-	+/-	+	-
497	-	+	+	-
498	-	-	-	-
4916	-	-	-	-
555	-	-	-	-
5510	-	+/-	-	-
5513	-	+	+++	-
564	-	-	-	-
566	-	-	-	-
581	-	-	-	-
788	-	-	-	-
902	-	-	-	-
9318	-	-	-	-
2914	-	+	-	-
417	-	+++	+++	-



D. Cyclodipeptides issus de la souche 497 : identification structurale et réseau moléculaire

1) Déréplication

Lorsque nous avons initié le travail de déréplication d'un extrait actif pour isoler des métabolites antimicrobiens, les évaluations biologiques des activités antagonistes des actinobactéries isolées et de leur extrait étaient en cours de réalisation. Pour des contraintes de temps il était nécessaire de sélectionner une souche candidate avant de connaître l'ensemble des résultats issus du criblage. Les 3 extraits de la souche 497 présentaient alors une forte inhibition contre les deux souches pathogènes *S. aureus* et *E. coli* et nous avons décidé de nous intéresser à cette souche. Notons également que les résultats du séquençage Sanger n'étaient pas connus et la souche 497 n'était donc pas identifiée. Cette actinobactérie a été isolée sur la cuticule de fourmis *Labidus sp.* prélevées sur le sentier du Loyola à Rémire-Montjoly en Janvier 2014. Dans un premier temps le fractionnement bioguidé de l'extrait (200 mg) sur colonne ouverte de silice C18 nous a permis d'obtenir 15 fractions dont 3 fractions se sont montrées actives contre *S. aureus* et *E. coli* (Figure 24 et Table 7). Nous avons ensuite effectué la purification d'une quantité plus importante d'extrait (1 g) par analyse chromatographique liquide semi-préparative à haute pression (HPLC) et nous avons tenté d'isoler les pics correspondants aux fractions actives. Cependant la séparation des métabolites d'intérêt s'est révélée délicate et un seul produit a pu être isolé avec une pureté suffisante pour être caractérisé (composé 1: 6,5 mg).

Table 7 : Résultats des CI_{50} des fractions F4, F5 et F6.

CI_{50} ($\mu\text{g/mL}$)	F4	F5	F6
<i>S. aureus</i>	$1,8 \pm 1,2$	$3,6 \pm 0,3$	$2,0 \pm 0,2$
<i>E. coli</i>	$8,1 \pm 0,6$	$34,8 \pm 0,1$	$18,4 \pm 0,2$

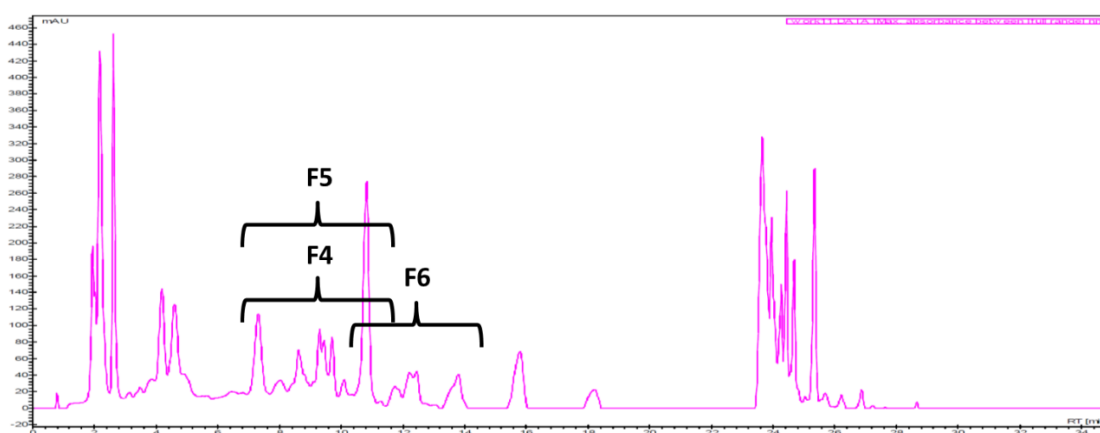


Figure 24 : Chromatogramme HPLC de l'extrait brut de la souche 497, indications des pics présents dans les fractions actives F4, F5 et F6.

2) Identification structurale

L'analyse par spectrométrie de masse à haute résolution du composé isolé ($[M+H]^+$ 245,1288) indique la formule brute $C_{14}H_{16}N_2O_2$. Les spectres RMN 1H , ^{13}C , COSY, gHSQC, HMBC indiquent la présence d'un groupement phényle monosubstitué, deux groupements méthines et 4 groupements méthylènes (Figure 25, Figure 26, Table 8). Le méthylène (3,15/38,21 ppm, $C_{10}H_2$) possède une corrélation C-H à longue distance avec deux carbones aromatiques (131,04, 137,32 ppm respectivement $C1'$ et $C2'$), un carbonyle d'une fonction amide (166,89 ppm, $C1$), et un méthine (4,44/57,49 ppm $C9H$), indiquant la présence d'une sous-unité de type phénylalanine. Par comparaison avec les données spectroscopiques de la littérature, l'autre unité est identifiée comme étant la proline. Au regard de ces résultats, le composé isolé est l'un des stéréoisomères d'une 2,5-dicétopipérazine la Cyclo(Pro-Phe), un peptide cyclique composé de deux acides aminés.

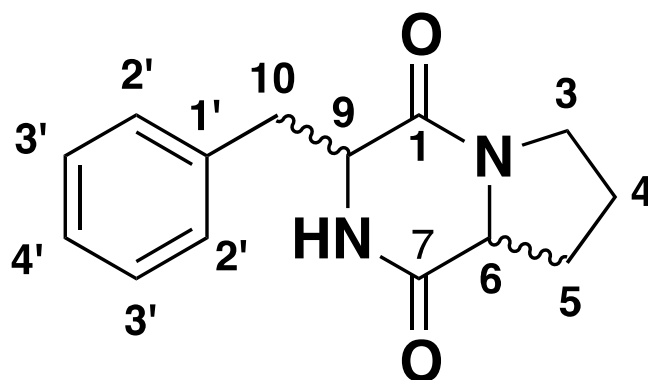


Figure 25 : Structure moléculaire de la dicétopipérazine (R, S)Pro-(R, S)Phe isolée.

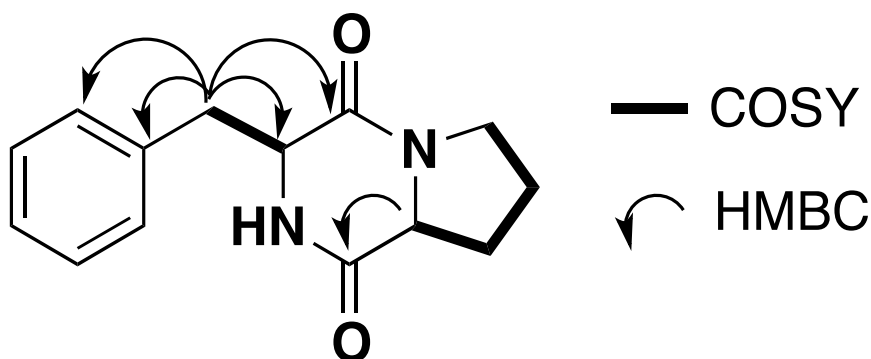


Figure 26 : Corrélation COSY et HMBC de la dicétopipérazine (R, S)Pro-(R, S)Phe isolée.

Table 8 : Table des résultats RMN de la dicétopipérazine (R, S)Pro-(R, S)Phe isolée.

N°	¹³ C	¹ H	COSY (1H--> 1H)	HMBC (1H--> ¹³ C)
1	166,89 (C)			
3	45,96 (CH ₂)	3,35-3,38 (1H, m) 3,53-3,56 (1H, m)	H-3b, H-4 H-3a, H-4	
4	22,77 (CH ₂)	1,78-1,81 (2H, m)	H-3a, H-3b, H-5a, H-5b	
5	29,38 (CH ₂)	1,20-1,27 (1H, m) 2,09-2,11 (1H, m)	H-5b, H-4, H-6 H-5a, H-4, H-6	
6	60,07 (CH)	4,09 (1H, dd, <i>J</i> =10,0 et 3,2 Hz)	H-5a, H-5b	C-7
7	170,9 (C)			
9	57,69 (CH)	4,44 (1H, m)	H-10	
10	38,21 (CH ₂)	3,17 (2H, d, <i>J</i> =4,89 Hz)	H-9	C-1, C-1', C-2', C-9
Ar				
1'	137,32 (C)			
2'	131,04 (CH)			
3'	129,44 (CH)	7,20-7,30 (5H, m)		
4'	128,07 (CH)			

La modélisation des spectres de dichroïsme circulaire correspondant aux quatre stéréoisomères a permis, via leur comparaison avec le spectre de dichroïsme circulaire de la molécule 1, de conclure que c'est l'énantiomère cyclo(*L*-Pro-*L*-Phe) qui a été isolé Figure 27 : Comparaison des spectres de dichroïsme circulaire expérimental et modélisé. Les données spectrales de spectroscopie RMN de ce dipeptide de la littérature comparée aux déplacements chimiques des protons ¹H des spectres RMN enregistrées indiquaient l'absence des stéréoisomères cyclo(*L*-Pro-*D*-Phe) et cyclo(*D*-Pro-*L*-Phe), mais c'est seulement grâce aux données de modélisation des spectres dichroïques qu'il est impossible de connaître lequel des énantiomères cyclo(*L*-Pro-*L*-Phe) ou cyclo(*D*-Pro-*D*-Phe) a été isolé.

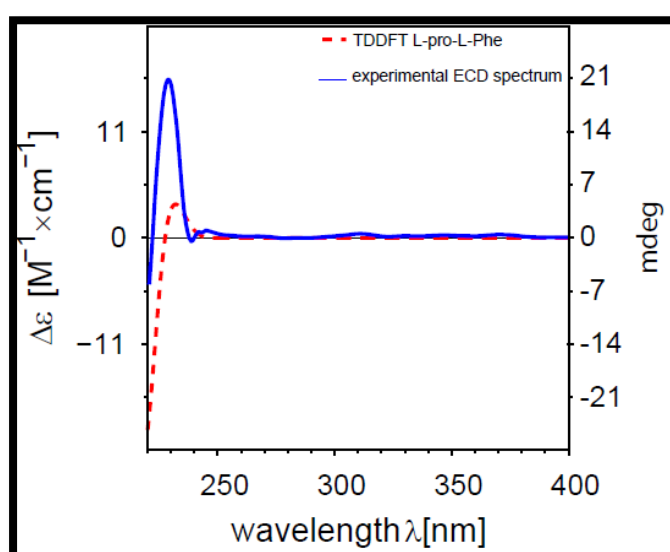


Figure 27 : Comparaison des spectres de dichroïsme circulaire expérimental et modélisé.

3) Les dicétopipérazines

L'analyse de la séquence ADN codant pour l'ARN 16S montre que l'espèce la plus proche de la souche 497 est *Streptomyces thioluteus* (base de données Eztaxon, pourcentage de similarité 99,28%, numéro d'accès : AJ781360). La 2,5-dicétopipérazine L-Pro-L-Phe isolée à partir de la souche 497 est la troisième molécule de cette famille identifiée dans la souche *Streptomyces thioluteus* (Gerber 1967). De plus, rappelons que de nombreuses 2,5-dicétopipérazines ont été isolées à partir d'autres espèces du genre *Streptomyces* (Ameur *et al.* 2004) (Pettit *et al.* 2006). Cette classe de molécule issue de la condensation de deux acides aminés est abondante dans la nature mais se retrouve également comme produit de dégradation de polypeptides dans l'industrie agroalimentaire (Borthwick 2012). Ces composés sont produits dans les règnes végétaux, animaux et microbiens et peuvent arborer divers groupements fonctionnels aboutissant à une grande diversité de molécules complexes. Ces molécules présentent des affinités avec de nombreux récepteurs impliqués dans des activités pharmacologiques notables : cytotoxique, antivirale, antibiotique, antifongique, neuroprotectrice, anxiolytique, anti-inflammatoire, antagoniste des récepteurs opioïdes, etc. (Figure 28) (Borthwick 2012).

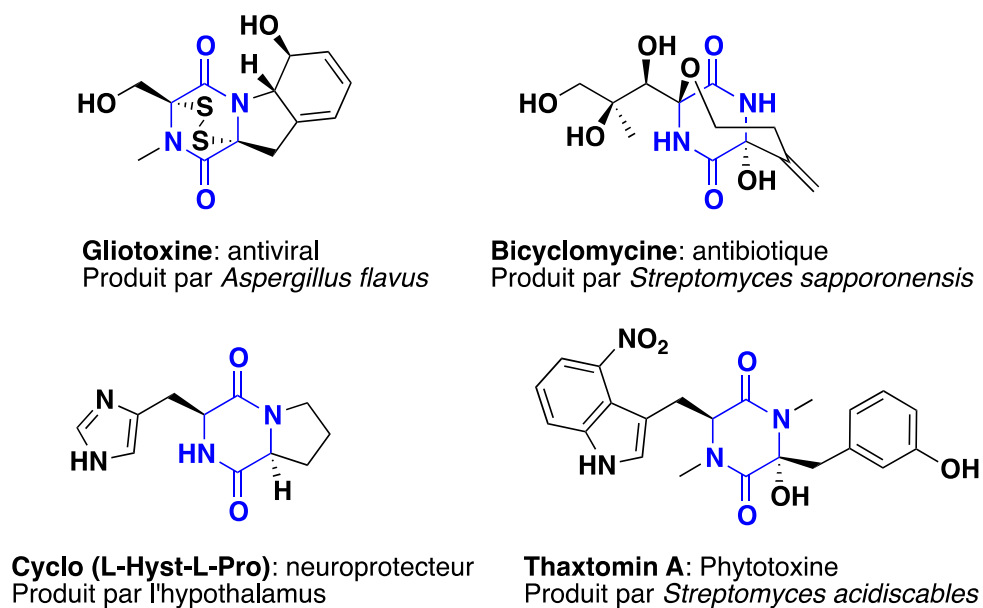


Figure 28 : Exemple de structure chimique de dicétopipérazines et de leurs activités biologiques respectives.

La structure rigide du cycle à 6 atomes des dicétopipérazines contenant deux carbones asymétriques (4 stéréoisomères possibles), peut mimer la conformation préférentielle d'un peptide sans pour autant posséder les propriétés indésirables des peptides (sensibilité à la protéolyse) ni les contraintes liées à la planarité des petites molécules pharmaceutiques. De plus cette structure est facilement synthétisable en chimie organique. Pour toutes ces raisons les 2,5-dicétopipérazines suscitent un intérêt croissant dans de nombreuses applications. On notera l'intérêt croissant en biologie chimique et en génie génétique pour explorer la diversité génétique des protéines assurant la biosynthèse des dipeptides. Récemment l'identification de 51 cyclo-dipeptide synthases (CDPs) d'origine microbienne suggère une grande diversité phylogénétique de ces enzymes et permet également de prédire la structure des dicétopipérazines qu'il reste à découvrir (Jacques *et al.* 2015).

Enfin dans un contexte d'écologie chimique il est intéressant de remarquer que plusieurs 2,5-dicétopipérazines ont été isolées à partir de la souche fongique *Tyridiomyces formicarum* cultivée par la fourmi attine *Cyphomyrmex minutus* (Wang *et al.* 1999). Ces résultats publiés en 1999 montraient pour la première fois que le champignon symbiotique des fourmis pouvait produire des molécules de défense contre l'invasion de souches fongiques pathogènes. Ainsi les 2,5-dicétopipérazines sont des produits naturels abondants qui sont utilisés comme régulateurs chimiques dans les interactions biologiques entre les fourmis et les micro-organismes.

Afin de mieux évaluer la biodiversité chimique des 2,5-dicétopipérazines de la souche 497 nous avons réalisé le réseau moléculaire de l'extrait de 497. Pour rappel, dans les réseaux moléculaires générés par spectrométrie de masse le nœud d'un réseau est formé lorsqu'un ion partage une similarité de ses fragments (score de cosine) supérieur à 0,7 avec au minimum un autre ion (cluster de 2 nœuds). L'ion correspondant à la dicétopipérazine isolée ici ($[M+H]^+$ 245,128 m/z) est dans un cluster, dans une sous partie du réseau, où d'autres dicétopipérazines se trouvent. Cette analyse nous permet d'identifier des structures connues et inconnues dans les bases de données en plus de la dicétopipérazine que nous avons isolé. On constate une diversité chimique de ces molécules dans l'extrait avec la présence de 8 nœuds sur le cluster (Figure 29). L'activité antimicrobienne observée est probablement le résultat de l'activité biologique cumulée ou synergique de ces composés.

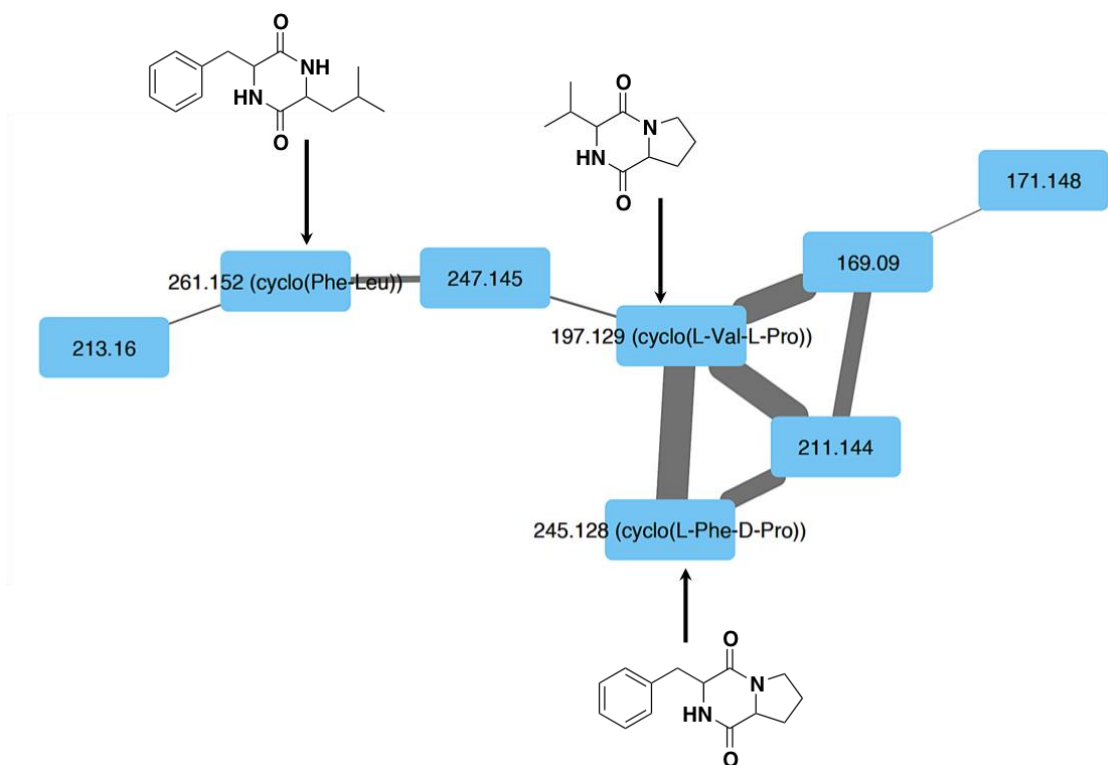


Figure 29 : Réseau moléculaire et structure chimique des 2,5 dicétopipérazines identifiées dans la souche 497.

4) Autres métabolites issus de la souche 497

Streptomyces thioluteus est une actinobactérie connue pour produire de nombreux métabolites mais dont le génome complet n'a pas encore été séquencé. L'auréothine (Figure 30) est un inhibiteur de l'oxydoréductase aux activités antifongiques, antibactériennes et cytotoxiques isolé pour la première fois dans les années 1950 dans l'extrait de culture de *Streptomyces thioluteus* (Hirata *et al.* 1954; Yoshimasa *et al.* 1961). Les travaux actuels dans ce domaine s'intéressent d'une part à la biosynthèse (Adermann *et al.* 2004) (Figure 30) et à la synthèse totale de l'auréothine (Henrot *et al.* 2016), notamment pour la production d'analogues d'intérêt pharmacologique. D'autre part, le groupement nitro (NO₂) est une fonction chimique rarement rencontrée dans la structure des produits naturels, et sa formation est étudiée pour comprendre le mécanisme d'oxydation par lequel l'enzyme AurF transforme l'acide *p*-aminobenzoïque (NH₂-Ar-COOH) en acide *p*-nitrobenzoïque (NO₂-Ar-COOH) (Li *et al.* 2010). Dans nos analyses par spectrométrie de masse nous avons détecté un ion ([M+H]⁺ 398,1573) de la même masse moléculaire que l'auréothine. De plus, par comparaison avec les données spectrales de fragmentation par MS² de la littérature, cet ion correspond effectivement à l'auréothine (Figure 31) (Shao *et al.* 2011). D'autre part, des ions correspondants à des métabolites de type phénazine et phénoxazinone (questiomycin A) ([M+H]⁺ 213,0642) ont été

détectés par spectrométrie de masse. Ces molécules avaient été précédemment identifiés dans la souche *Streptomyces thioluteus* (Gerber 1967).

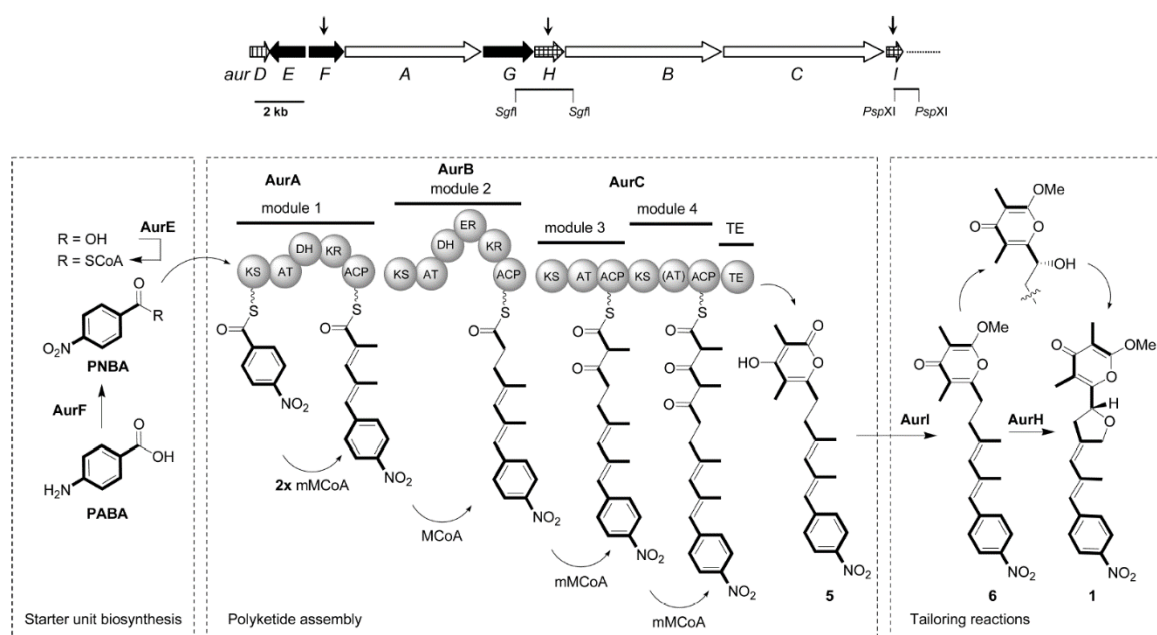


Figure 30 : Biosynthèse de l'auréothine (Adermann *et al.* 2004).

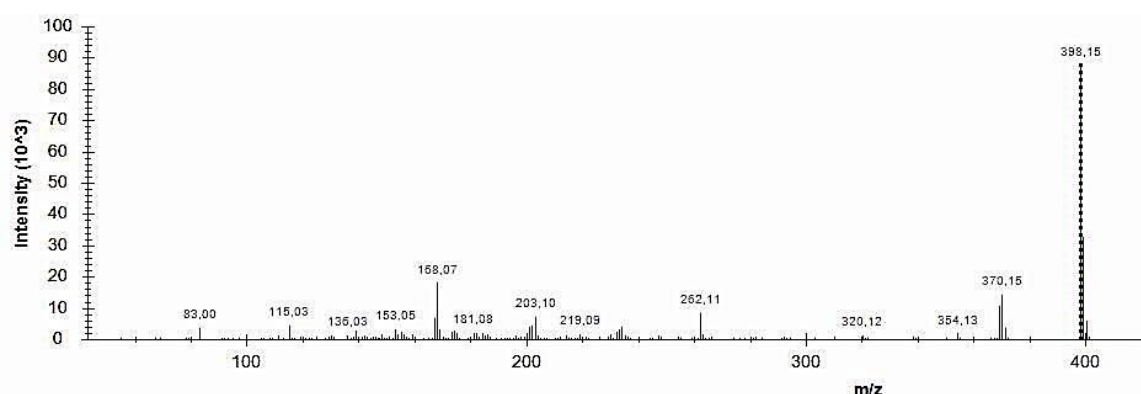


Figure 31 : Spectre expérimental MS² de l'auréothine. (Auréothine standard : MS¹ :398.15 [M+H]⁺, MS²: 370,1 262,0 203,0) (Shao *et al.* 2011).

E. Induction de la production des métabolites d'actinobactéries

1) Présentation

Afin de d'augmenter la production de métabolites antibiotiques de manière qualitative et/ou quantitative, nous avons fait le choix de stresser les souches d'actinobactéries par d'autres souches d'actinobactéries. Cette approche est tout à fait pertinente pour la prospection de nouvelles molécules antibiotiques au travers des interactions en place au sein de la communauté bactérienne cuticulaire. Nous avons choisi des souches d'actinobactéries isolées à partir de la même colonie d'*Ectatomma tuberculatum* prélevé en janvier 2014 à Cayenne. A partir de six

actinobactéries différentes, (souches 581, 5510, 564, 555, 566, 5513), nous avons réalisé 6 associations de souches (en duplicat pour chacune) (Figure 32) en nous basant sur les raisons suivantes :

- Toutes les associations sont réalisées à partir d'actinobactéries issues de la même colonie de fourmi. En effet, des souches d'une même fourmi pourraient avoir déjà développées des mécanismes de production de métabolites secondaires entre elles pour communiquer, se défendre et coloniser la surface sur la cuticule.
- L'originalité de ces 6 cocultures sont des associations par deux (CoC4, CoC5, CoC6) trois (CoC2, CoC3) et ce jusqu'à six souches (CoC1). Les cocultures décrites dans la littérature sont le plus souvent constituées de seulement deux micro-organismes, et mettre en jeux plus de deux souches pourrait donner lieu à des cocktails chimiques plus diversifiés.
- Les associations sont principalement construites autour d'une souche déjà active en monoculture (5513 et 5510) afin d'identifier une augmentation de l'activité antibiotique attendue.
- Certaines souches non actives en monoculture mais issues de la même colonie sont mises en contact tout de même, pour déterminer si l'apparition d'une activité antibiotique existe en coculture (CoC5).

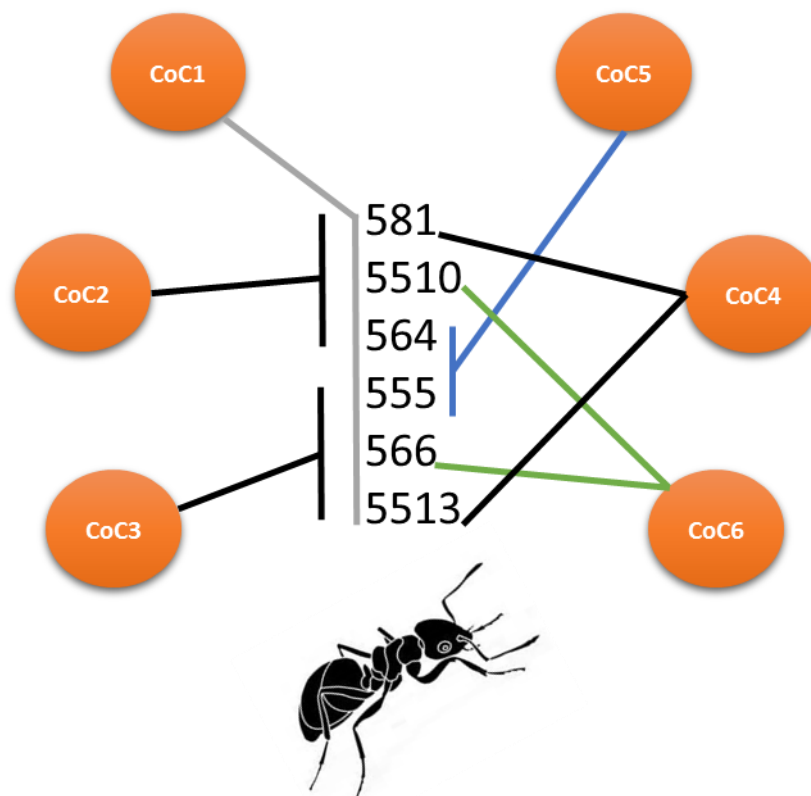


Figure 32 : Composition des six assemblages de cocultures, avec les codes des six souches d'actinobactéries au centre.

2) *Activité antibiotique*

Chaque extrait purifié est testé de manière identique dans 3 puits à 400 µg/mL contre les quatre souches bactériennes : *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus* et *Enterococcus faecalis*. Les résultats des inhibitions sont présentés dans la Table 9 ci-dessous et utilisent la même classification que précédemment.

Les pourcentages d'inhibition des extraits de coculture sont généralement supérieurs à ceux des extraits des souches en monoculture et inhibent un spectre plus large de souches pathogènes ce qui justifie la pertinence de cette approche pour de la prospection. Les extraits de la coculture avec les 6 souches (CoC1) inhibent *E. faecalis* et *P. aeruginosa* comme la souche 5513 en monoculture et inhibent *E. coli* et *S. aureus* alors qu'aucune des 6 souches prises séparément ne les inhibaient en monoculture. Parmi les expériences à 3 souches, la coculture CoC2 présente une augmentation de l'activité antibiotique pour *E. faecalis* et une nouvelle inhibition contre *P. aeruginosa*. La CoC3 inhibe les souches pathogènes testées de manière identique à la souche 5513 la composant avec une nouvelle inhibition (faible) contre *E. coli*. Les cocultures à 2 souches permettent de confirmer que les souches 555 et 564 non actives ne

le sont pas non plus en coculture (CoC5). Elles permettent également de confirmer que la souche la plus active 5513 en coculture avec 581 seulement (CoC4) donne des extraits autant actifs qu'en monoculture. Tandis que 5510 en mélange avec 566 présente des extraits inhibant plus fortement *E. faecalis* et présente une inhibition de *P. aeruginosa* non observée en monoculture.

Table 9 : Pourcentage d'inhibition des extraits purifiés des 6 souches d'actinobactéries et des 6 cocultures contre 4 souches tests pathogènes, inhibition totale = +++ (rouge), inhibition forte entre 50% et 100% = + (orange) et inhibition faible entre 15% et 50% = +/- (jaune) et absence d'inhibition = - (blanc).

Code souche	Staphylococcus aureus	Enterococcus faecalis	Pseudomonas aeruginosa	Escherichia coli
555	-	-	-	-
5510	-	+/-	-	-
5513	-	+	+++	-
564	-	-	-	-
566	-	-	-	-
581	-	-	-	-
CoC1	+/-	+	+++	+
CoC2	-	+	+	-
CoC3	-	+	+++	+/-
CoC4	-	+	+++	-
CoC5	-	-	-	-
CoC6	-	+	+	-

Pour illustrer les différences chimiques observées entre les monocultures et les cocultures qui pourraient expliquer l'augmentation d'activité, le profil chromatographique de la coculture 4 (mise en commun des souches 581 et 5513) est présenté en Figure 33. Le profil chromatographique de la CoC4 ressemble principalement à la souche 5513. Cependant, dans le cas de CoC4 à concentration égale des extraits (1 mg/mL) et pour une même quantité d'extrait injecté (12 µL) l'intensité des pics majoritaires en HPLC (détection UV) est supérieure aux intensités de ces mêmes pics de la monoculture. La souche 581 stimule donc la production des métabolites de la souche 5513. Dans ce cas, le profil chimique de CoC4 présente la même complexité chimique que le profil issu de la somme des deux monocultures. Nous avons ici l'exemple d'une augmentation de la quantité de métabolites. Dans la partie suivante nous chercherons à identifier les métabolites nouvellement produits dans les cocultures.

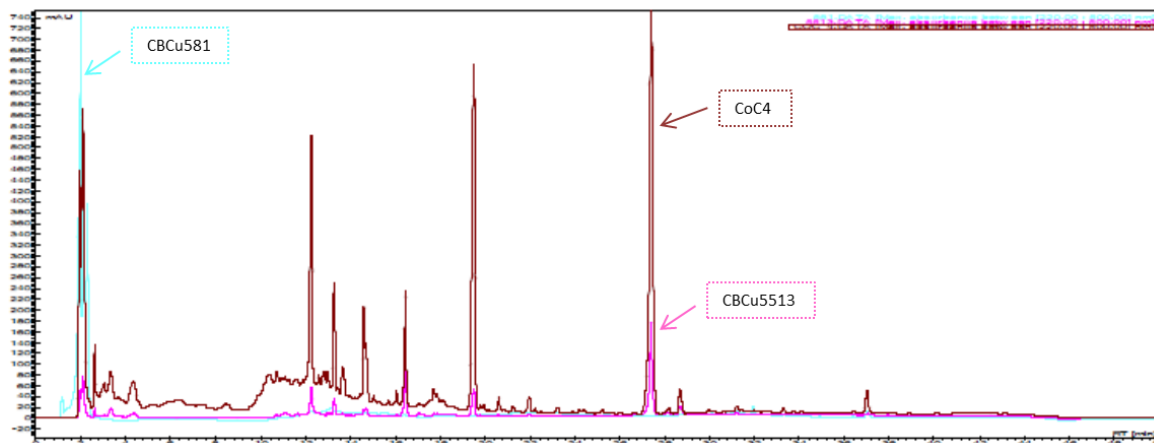


Figure 33 : Chromatogrammes HPLC UV superposés de la CoC4 et des deux monocultures 581 et 5513.

3) *Augmentation de l'activité de CoC1 et CoC2*

Comment expliquer l'augmentation d'activité antibiotique de la coculture CoC1 à six souches et celle de la coculture CoC2 à trois souches ?

Dans les réseaux moléculaires générés par spectrométrie de masse, le nœud d'un réseau est formé lorsqu'un ion partage une similarité de ses fragments (score de cosinus) supérieur à 0,7 avec au minimum un autre ion (cluster de 2 nœuds). Ainsi, toutes les molécules des monocultures et des cocultures ne sont pas représentées sur le réseau moléculaire, d'autant plus que toutes les molécules ne sont pas ionisables. Néanmoins, ces limitations sont les mêmes pour tous les extraits injectés, ce qui permet de comparer le pourcentage d'ions strictement nouveaux dans les différentes cocultures.

Parmi les extraits des cocultures CoC1 et CoC2 qui nous intéressent pour leur augmentation d'activité antibiotique, 34,1% des ions de CoC1 ne sont partagés par aucune des six monocultures et 33,2% des ions de CoC2 ne sont présents dans aucune des trois monocultures. Il est intéressant de noter que nous avons obtenu le même pourcentage d'ions nouvellement formés dans les deux cocultures indépendamment du nombre de souches mises en coculture. Ces résultats préliminaires ne permettent pas d'avancer que cette tendance est vérifiée dans tous les cas mais indique la pertinence d'utiliser les réseaux moléculaires pour faciliter l'identification des ions nouveaux recherchés.

Les réseaux moléculaires des figures ci-dessous (Figure 34 et Figure 35) permettent de visualiser la présence de métabolites produits en coculture. Ainsi chaque nœud est un point dont la couleur représente la présence de l'ion dans la ou les monocultures (couleur grise) ou dans

la coculture (couleur rose ou bleue). Un nœud totalement rose est un ion présent uniquement dans la coculture. Les souches en monoculture sont représentées dans les nœuds en gris et les nuances de gris déterminent les différentes souches. De cette manière il est possible de visualiser qualitativement et quantitativement la présence des ions dans les monocultures et dans les cocultures.

Les réseaux moléculaires générés à partir des données de spectrométrie MS/MS pour les deux cocultures CoC1 et CoC2 montrent la présence de nombreux clusters composés de 2 à plus de 50 ions. Au sein d'un cluster, que ce soit dans le réseau moléculaire de CoC1 (Figure 34) et CoC2 (Figure 35), on identifie des ions en plus grande abondance en coculture et on constate également un certain nombre d'ions présents uniquement dans les cocultures.

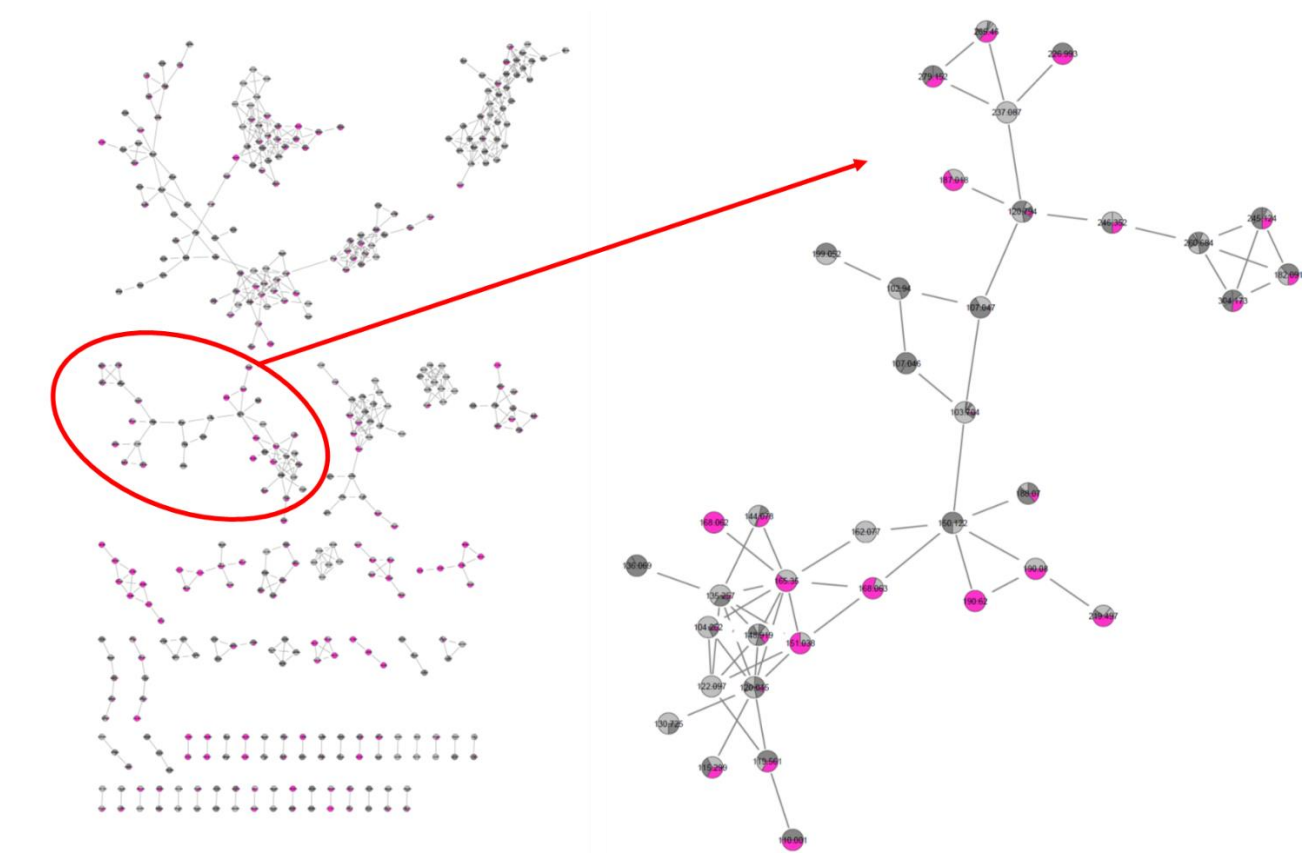


Figure 34 : Réseaux moléculaires des six souches en monoculture (gris) et de la coculture CoC1 (rose).

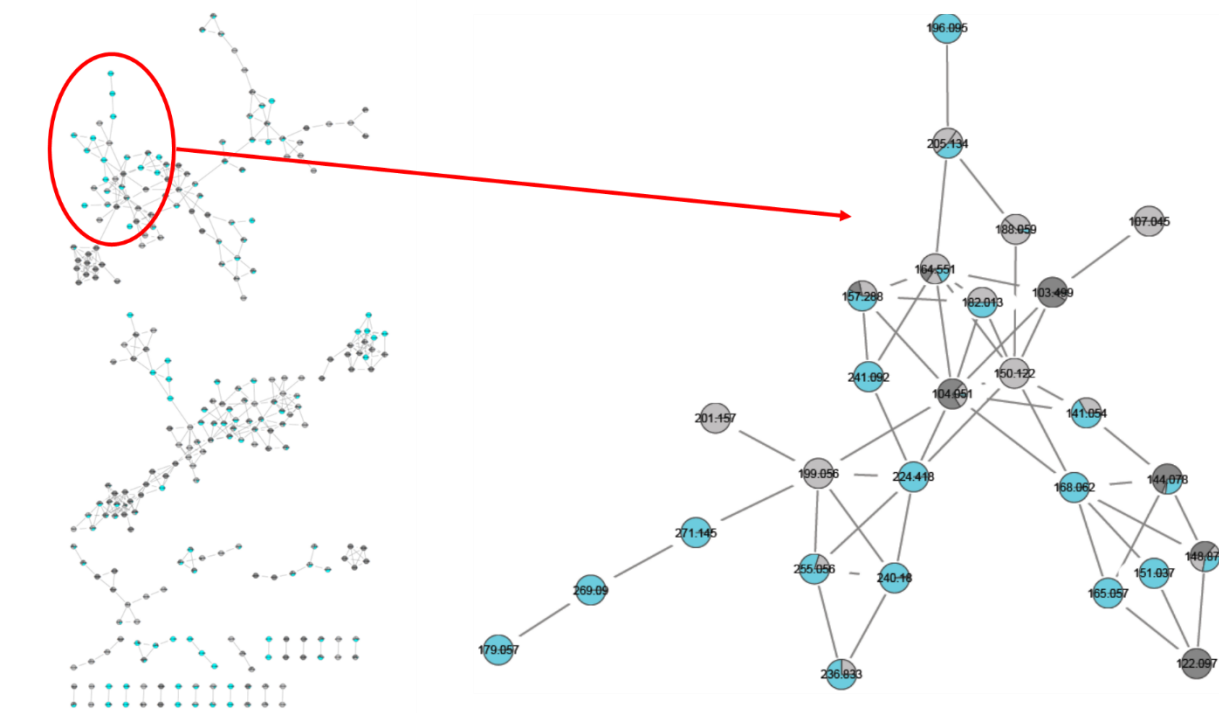


Figure 35 : Réseaux moléculaires des trois souches en monoculture (gris) et de la coculture CoC2 (bleu).

Cependant, l'identification des ions est limitée par la comparaison des valeurs de m/z des fragments avec celles des bases de données disponibles dans GNPS. Dans notre cas, nous avons identifié dans les réseaux moléculaires des 6 cocultures le cluster représentant les ions de 2,5-dicétopipérazine (non présenté). Il n'est cependant pas plus informatif que celui obtenu précédemment en monoculture (Figure 29), de plus il ne présente pas de composé nouvellement formé dans les cocultures.

Nous avons également généré les réseaux moléculaires à partir de l'ensemble des données issues des mono- et des cocultures. L'un des clusters de ces réseaux moléculaires possède deux ions identifiés comme étant probablement l'Antimycine A2 ($[M+H]^+$ 535,264 m/z) et l'Antimycine A3 ($[M+H]^+$ 521,233 m/z) (Figure 36). Les Antimycines sont des composés produits par le genre *Streptomyces* issus d'une voie de biosynthèse hybride de polyketide synthase (PKS) et peptide synthase non-ribosomique (NRPS) (Figure 37) (Seipke *et al.* 2011). Ces composés sont très toxiques car ils inhibent la chaîne respiratoire mitochondriale. Par ailleurs, ces composés ont été isolés à partir de différentes souches de *Streptomyces* prélevées sur la cuticule de fourmis champignonnistes (*Acromyrmex echinator*, *Acromyrmex octospinosus*, *Acromyrmex volcanus*) (Schoenian *et al.* 2011). Par conséquent, il n'est pas surprenant de mettre en évidence ces composés et il est très intéressant de noter que certains ions comme $[M+H]^+$ 535,264 m/z (Antimycin A2) et $[M+H]^+$ 549,28 m/z (non identifié) sont

produits uniquement dans la coculture CoC2. De nouveau ces résultats montrent la pertinence de combiner les stratégies de coculture et l'analyse des réseaux moléculaires pour à la fois augmenter quantitativement et qualitativement la production des métabolites, mais également pour visualiser et tenter d'identifier en peu de temps de potentielles molécules d'intérêt. Cette approche bioinspirée permet également une identification plus exhaustive des métabolites produits sur la cuticule des fourmis et renforce l'hypothèse du rôle bénéfique des communautés bactériennes dans la défense chimique de ces insectes sociaux

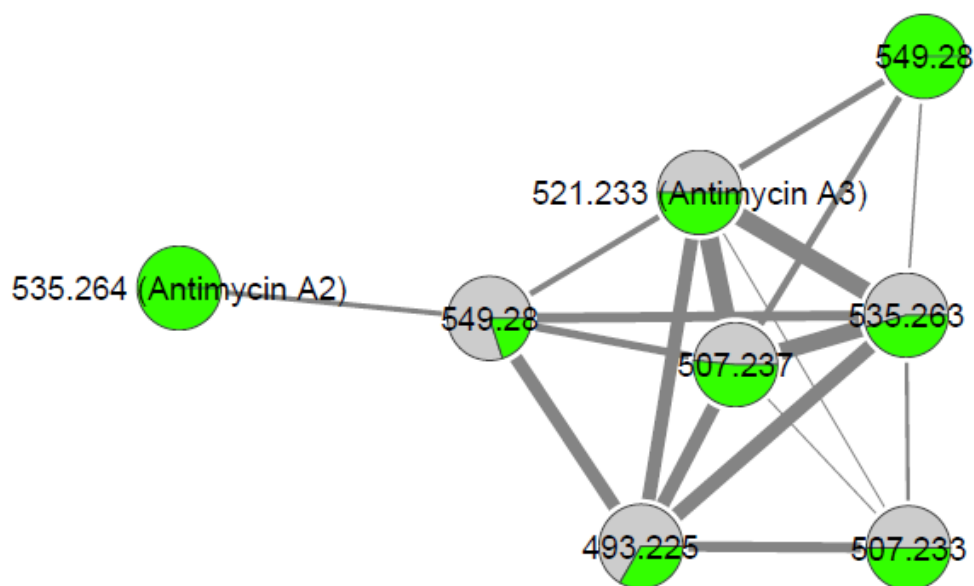
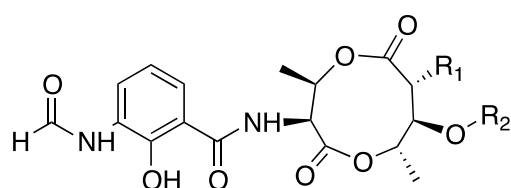


Figure 36 : Réseaux moléculaires des six souches en monoculture (gris) et des six cocultures CoC1, CoC2, CoC3, CoC4, CoC5, CoC6 (vert).



Antimycine A1: R₁ = C₆H₁₃; R₂ = COC₄H₉
 Antimycine A2: R₁ = C₆H₁₃; R₂ = COC₃H₇
 Antimycine A3: R₁ = C₄H₉; R₂ = COC₄H₉
 Antimycine A4: R₁ = C₄H₉; R₂ = COC₃H₇

Figure 37 : Structures chimiques des Antimycine A1 à A4.

Chapitre 2. Accéder à la diversité du microbiote bactérien cuticulaire.

I. Introduction

Le premier chapitre présente l'isolement de souches d'actinobactéries à partir de la cuticule des différentes espèces de fourmis (mais excluant les fourmis attines déjà très étudiées), et le pouvoir antibiotique de ces souches ainsi que la production de nouveaux métabolites dans les expériences de coculture. Ces éléments constituent des arguments en faveur d'un microbiote bénéfique sur la cuticule des fourmis, mais les résultats doivent être interprétés avec prudence. La présence d'actinobactéries sur la cuticule des fourmis ne signifie pas que ces souches, capables de produire des cocktails antibiotiques, sont effectivement présentes dans une relation symbiotique mutualiste avec leurs hôtes. Le microbiote de la cuticule des fourmis, bénéfique ou non, est éminemment plus diversifié que les quelques actinobactéries cultivables isolées. Grâce aux technologies haut débit et au séquençage du gène codant pour l'ARNr 16S, il est aujourd'hui possible d'accéder à la diversité du microbiote cuticulaire bactérien des fourmis afin de i) révéler le ou les rôles de ces bactéries sur la cuticule et ii) de comprendre comment s'effectue l'assemblage des communautés microbiennes.

La relation symbiotique mutualiste très étudiée entre des actinobactéries cuticulaires et les fourmis attines est une hypothèse controversée notamment concernant l'acquisition de ces souches supputées bénéfiques (Mueller 2012). Le recrutement des souches bénéfiques s'effectue-t-il dans l'environnement ou la transmission des souches bénéfiques est-elle verticale ? L'engouement pour l'étude de cette interaction bénéfique, détaillée précédemment, s'est principalement focalisée sur deux points : la valorisation de ces micro-organismes mutualistes en *drug discovery* et l'exploration du rôle défensif du microbiote cuticulaire. Nous notons qu'aucune étude utilisant le séquençage haut débit pour décrire la diversité des communautés bactériennes cuticulaires des fourmis attines n'a été menée jusqu'à présent. Nous pensons qu'une meilleure compréhension des rôles et des mécanismes d'assemblage de ces communautés passe d'abord par la caractérisation de la diversité de ce microbiote. La valorisation des métabolites secondaires de ces micro-organismes, la compréhension des rôles du microbiote cuticulaire, de ses mécanismes d'assemblage ainsi que la caractérisation de la diversité du microbiote sont autant de champs interdisciplinaires nécessaires à une meilleure compréhension de l'interaction intime entre les fourmis attines et les actinobactéries mais plus globalement aux interactions entre le microbiote cuticulaire et les fourmis non-attines.

La métagénomique, grâce à l'arrivée des séquenceurs à haut débit utilise les ADN présents dans l'échantillon pour accéder à la description de la diversité, ce qui est plus efficace

par rapport à l'identification des micro-organismes cultivables de l'échantillon. Toutefois, ces techniques comportent des limites. Une quantité suffisante d'ADN extrait est nécessaire, ainsi les petits échantillons tels que les micro-organismes issus du lavage d'une fourmi sont problématiques et fortement sujet aux contaminations extérieures. La littérature disponible sur l'utilisation de ces techniques d'extraction d'ADN jusqu'aux traitements bioinformatiques est récente, et à notre connaissance, ces techniques n'ont pas fait l'objet d'études méthodologiques pour ce type d'échantillon.

Ces derniers éléments, ainsi que de nombreux échecs d'extraction des ADN de nos échantillons, nous ont conduits à mener une étude comparative de 4 protocoles différents d'extraction d'ADN pour accéder à la diversité des bactéries cuticulaires d'un unique individu fourmi. L'objectif de ce travail est de maximiser l'étape d'extraction d'ADN pour évaluer la diversité bactérienne non-cultivable. Nous évaluons ici la richesse et la composition du microbiote cuticulaire des échantillons extraits par les quatre méthodes, ainsi que l'effet du nombre d'individu ou la taille des individus. Ce chapitre 2 correspond à l'article tel qu'il a été soumis à Molecular Ecology Ressources en décembre 2016, et il est par conséquent rédigé en anglais dans ce manuscrit de thèse.



Figure 38 : Photos des fourmis *Pseudomyrmex penetrator* (à gauche) (auteur : Caroline Birrer) *Atta cephalotes* (à droite) (auteur : Alex Wild <http://www.alexanderwild.com>).

II. Comparative analysis of DNA extraction methods to study the body surface microbiota of insects: a case study with ant cuticular bacteria

Authorship

Caroline Birer^{a*}, Niklas Tysklind^b, Lucie Zinger^{c,d}, Christophe Duplais^a

Author contributions

All authors contributed to the conception of the study. C.B. and C.D. performed the sampling, C.B. carried out the molecular work, L.Z. performed the bioinformatics analyses and C.B. and L.Z. conducted the statistical analyses. All authors have significantly contributed to the writing of the manuscript.

A. Abstract

High-throughput sequencing of the 16S rRNA gene has considerably helped revealing the essential role of bacteria living on insect cuticles in the ecophysiology and behavior of their hosts. However, our understanding of host-cuticular microbiota feedbacks remains hampered by the difficulties to working with low bacterial DNA quantities as in individual insect cuticle samples, which are more prone to molecular biases and contaminations.

Herein, we conducted a methodological benchmark on the cuticular bacterial loads retrieved from two Neotropical ant species of different body size and ecology: *Atta cephalotes* (~15 mm) and *Pseudomyrmex penetrator* (~5 mm). We evaluated the richness and composition of the cuticular microbiota, as well as the amount of biases and contamination produced by four DNA extraction protocols. We also addressed how bacterial communities characteristics would be affected by the number of individuals or individual body size used for DNA extraction.

Most extraction methods yielded similar results in term of bacterial diversity and composition for *A. cephalotes*. In contrast, greater amounts of artifactual sequences and contaminations, as well as noticeable differences in bacterial communities characteristics were observed between the extraction methods for *P. penetrator*. We also found that large and small *A. cephalotes* individuals harbor different bacterial communities. Our benchmark hence suggests that cuticular microbiota of single insect individuals can be reliably retrieved provided that blank controls, appropriate data cleaning, and standardization of individual life stage are considered in the experiment.

B. Introduction

The development of high-throughput methods in molecular ecology has enabled exploring the microbial world with an unprecedented level of resolution and helped reveal their fundamental roles in the functioning of all ecosystems (Moran & Sloan 2015). Such interdependence and potential coevolution between host and microbiota, calls for a better understanding of the nature of the associated microbiota, and disentangling their part in the evolutionary success of their hosts. While the gut microbiomes is increasingly studied (Cogen *et al.* 2008; Ley *et al.* 2008), the microbiome of the skin, or more generally of body surfaces, remains comparatively largely under-investigated.

However, the skin is one of the largest organs of animals, and constitutes their first physical and biological barrier against pathogens and toxic agents (Vallet-Gely *et al.* 2008; Grice & Segre 2011). Animal body surfaces harbour diverse microbial communities in both vertebrates and invertebrates, including in humans (Oh *et al.* 2014), amphibians (Kueneman *et al.* 2014), fish (Boutin *et al.* 2014), and insects (Suen *et al.* 2010). These body surface microbial communities are assembled according to environmental exposure (Sylvain *et al.* 2016), host species specificity (Larsen *et al.* 2013), and even host genotype regulation (Boutin *et al.* 2014). Body-surface microbial communities are thought to be involved in host defence mechanisms through physical blocking of pathogens by biofilm production, competition with pathogens for nutrient access, or biosynthesis of antimicrobial metabolites (Brownlie & Johnson 2009). These interactions can be of such significance that some insect species have evolved particular cuticular structures that favour the colonization of beneficial microbial communities (Douglas 2015). Notable examples of such structures include mycangia, a cuticular invagination housing fungi in ambrosia beetles (Coleoptera) (Hulcr *et al.* 2012), cuticle-lined glandular reservoirs in antennal segments of solitary digger wasps (Hymenoptera) (Kaltenpoth *et al.* 2012), and crypts and foveae in fungus-growing ants (Hymenoptera) (Currie *et al.* 2006). These structures usually harbour high levels of Actinobacteria, a bacterial group well known for its antimicrobial activities (Oh *et al.* 2009a).

The above illustrates the strong interplay between insect cuticle microbiota and their host. However, the evolutionary and ecological mechanisms shaping cuticle microbiota remain poorly understood (Scheuring & Yu 2012). A methodological challenge prevents detailed description of insect cuticular microbiota: bacterial DNA extraction protocol, which involves washing one small individual in a detergent solution to detach low amounts of bacterial cells tightly fixed on the cuticle. Choosing an appropriate DNA extraction method here is crucial as they differ in the amount of DNA lost during the elution step, in their efficiency to remove PCR

inhibitors contained in cuticles, bacteria, or washing solutions. Working with minute amounts of DNA also increases the risk of contamination (van Oorschot *et al.* 2010); DNA contaminants can be introduced during the processing of samples, including through consumables such as plastic tubes, handling equipment, lab reagents (Ishak *et al.* 2011; Salter *et al.* 2014), and commercial DNA extraction kits (Mohammadi *et al.* 2005; Gefrides *et al.* 2010; Goodrich *et al.* 2014; Salter *et al.* 2014). The concentration of non-target DNA in the final DNA extract can equal or surpass the concentration of targeted DNA and lead to grossly erroneous conclusions, especially when using highly sensitive techniques such as PCR and high-throughput sequencing (Knapp & Hofreiter 2010; Thomsen & Willerslev 2015). Furthermore, the sample multiplexing strategy commonly used in metabarcoding studies is prone to false positives. These arise from switches of DNA labels (hereafter tags) used to recognize the sample origin of a given sequence during library preparation or sequencing (termed "tag-switching") and causes incorrect assignment of amplicons to samples (Esling *et al.* 2015; Schnell *et al.* 2015). To circumvent these drawbacks, one common practice is to perform DNA extractions on pools several insect individuals (from 5 to 10) to successfully assess the bacterial diversity on insect's cuticles (Seipke *et al.* 2013). However, pools of individuals preclude determining the extent of variation of microbial communities across individuals.

These considerations emphasize the need to thoroughly evaluate DNA extraction protocols when working with such challenging material if we are to accurately unveil assembly mechanisms of insect cuticle microbial communities. Herein, we provide a comprehensive assessment of different DNA extraction protocols in their ability to minimize molecular contamination and to retrieve the richness, composition, structure of cuticular bacterial communities associated with two Neotropical ant species: *Atta cephalotes* and *Pseudomyrmex penetrator*. The former is a fungus-growing leaf-cutter ant that nests underground and exhibits extreme body size and labour polymorphism among workers (Wetterer 1999). The latter is an obligate mutualist of the myrmecophyte *Tachigali spp* nesting inside domatia at the base of the leaves. The average size of *A. cephalotes* varies from 2 mm to 15 mm, while that of *P. penetrator* workers is of ca. 5 mm. Such disparity in body size between the two species, as a surrogate for disparity in their body surface area, could imply significant differences in their bacterial loads and associated amounts of DNA. Therefore, we also included in our benchmark samples made of *A. cephalotes* individuals of 5 mm versus 15 mm in order to distinguish between the effects of body size and species. Finally, we also assessed the feasibility of extracting DNA from the cuticular microbiome from single individuals as compared to those retrieved from pools of 10 individuals.

C. Material and Methods

1) *Taxon sampling*

Fungus-growing *A. cephalotes* workers of two different sizes and performing different tasks (*minor* parasol riders: 5 mm and *major* soldiers: 15 mm) were collected outside their nest on colony trails at Sentier des Trois Rois (Latitude: 4° 56' 339'' N, Longitude: 52° 32' 051'' W); while myrmecophytic *P. penetrator* workers (5 mm) were collected from *Tachigali paniculata* Aubl. trees at La Montagne des Singes (Latitude: 5° 4' 435'' N, Longitude: 52° 41' 957'' W). Both sites in French Guiana were sampled in June 2015. For each ant species, at least 200 individuals were collected within the same colony to avoid introducing inter-colony variation that could mask methodological effects. Individuals were transferred to sterile plastic tubes with sterilized tweezers that were flamed and rinsed in ethanol between individual collections and were preserved the same day at -20°C. Myrmecologists F. Petitclair and Dr. A. Touchard confirmed species identifications.

2) *Overview of DNA extraction protocols*

Our comparative study included four DNA extraction methods or kits: the Qiagen QIAamp DNA mini kit (QIAamp; Qiagen, Carlsbad, CA, US), the Qiagen DNeasy blood and tissue kit (Dneasy; Qiagen, Carlsbad, CA, US), a modified phenol-chloroform protocol (PC) (Sambrook *et al.* 1989), and the MO BIO UltraClean tissues and cells kit (MO BIO; Mo Bio Laboratories, Hilden, Germany). To evaluate their performance, we applied the methodology depicted in Figure 39. First, thirteen individuals of *P. penetrator* workers and *A. cephalotes* workers were washed separately for each extraction method. To ascertain the sources of DNA contamination in each extraction method, two types of negative controls were included: a tweezers control and an extraction control. The tweezers control measures the contamination introduced by the manipulation of samples during the washing step. The extraction control measures the contamination introduced by the DNA extraction protocol.

To test the effect of the starting material (i.e. pooling or not individuals), four groups of 10 individual workers of each *P. penetrator* (5mm) and *A. cephalotes* majors (15mm) were pooled. To disentangle the effect of species vs. body surface, four additional groups of 10 individual *A. cephalotes* minor workers (5mm) were pooled. Total DNA of pools was extracted with QIAamp DNA mini kit only.

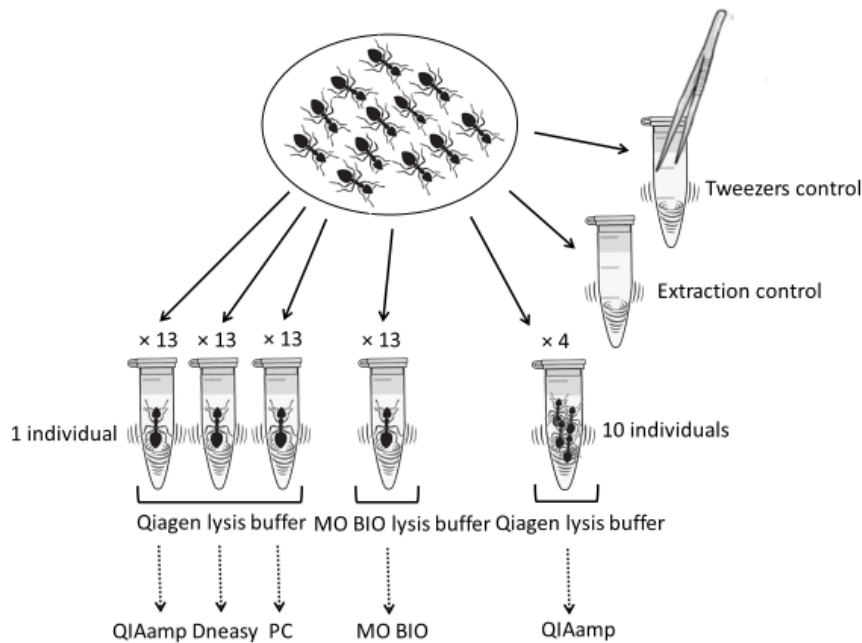


Figure 39 : Experimental design for evaluating bacterial cuticular microbiota using different DNA extraction protocols. The design was applied to two ant species are *A. cephalotes* and *P. penetrator*, and the extraction protocols tested are QIAamp: Qiagen QIAamp DNA mini kit, Dneasy: Qiagen DNeasy blood and tissue kit, PC: Phenol-chloroform, MO BIO: MO BIO UltraClean tissues and cells.

3) *Ant washing and DNA extractions*

Prior DNA extraction, ants were put 15 minutes at -20 °C to avoid a stress response caused by the wash that may lead to faecal evacuation. All the experiment was carried out in a sterile environment. The DNA extraction methods tested here followed classical protocols, of which details are given in Appendix S1. Note that QIAamp and Dneasy kits differ in the pore size of the silica-membrane. QIAamp has smaller pores, which would limit DNA losses.

The complete experimental design is composed of 138 extractions: 52 individual cuticular DNA extracts for each ant species *P. penetrator* and *A. cephalotes*; 4 and 8 pools of washes from 10 individuals of *P. penetrator* and *A. cephalotes* respectively; as well as 11 DNA extracts of each tweezers and extraction controls.

4) *DNA quantification*

Total DNA concentrations of all samples were measured on a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, US) with the dsDNA High Sensitivity Assay Kit (Life Technologies Corp., Oregon, US) using 5 µL of DNA extract.

5) *PCR amplification and sequencing*

We used the V5-V6 regions of the 16S rRNA gene to assess bacterial community diversity and composition. The universal primer pair used were BACTB-F

(GGATTAGATACCCTGGTAGT) and BACTB-R (CACGACACGAGCTGACG), which amplifies a fragment of about 295 bp in length (Fliegerova *et al.* 2014). Both forward and reverse primers were labelled with a combination of two different 8 nt tags, so samples could be multiplexed into a single sequencing lane while retaining the ability to reassign sequences to samples. PCR amplifications were performed in duplicate for each sample to evaluate amplification consistency. The PCR mixtures consisted of 1 µL of extracted DNA, 0.5 µM of each primer (Eurogentec, Angers, France), 31.25 µM of each dNTP, 2.5 mM of MgCl₂, 0.05 U/µL AmpliTaq Gold Polymerase (Applied Biosystems, Foster City, CA, USA), and 1x AmpliTaq Gold Buffer, and UHQ DNA-free water to reach a final volume of 20 µL. The PCR amplification consisted of an initial denaturation step at 95°C for 10 min, followed by 30 cycles of 95°C for 30 s, 57°C for 30 s, and 72°C for 30 s, and a final extension step at 72 °C for 7 min. In addition to tweezer and extraction controls, PCR controls were introduced to evaluate the sources of DNA contamination incorporated at the PCR step. All PCRs, including controls, were pooled, purified and the library was prepared with the Illumina TruSeq Nano kit for Illumina sequencing. The library, hence comprised a total of 279 PCRs, was then sequenced on one lane of Illumina MiSeq platform using paired-end 2x250 bp technology (Genotoul, Toulouse, France).

6) Sequence analysis

We obtained a total numbers of 2×9,308,468 reads which were subjected to a standard data curation process mainly based on the OBITools package (Boyer *et al.* 2016). First, paired-end reads were assembled into extended single sequences, taking into account the read overlapping quality, with the FLASH algorithm (Magoč & Salzberg 2011). Second, sequences were assigned to their respective sample by allowing 0 errors in tags and a maximum of two errors on primers. Sequences that did not fulfil these criteria were discarded, as well as those that are most likely sequencing errors (i.e. containing ambiguous nucleotides or shorter than 100 bp). Finally, the remaining sequences were dereplicated and clustered into OTUs (Operational Taxonomic Units) with Sumacust (Mercier 2013) with a sequence similarity threshold at 97%. The most abundant sequence of each OTU was considered as the OTU representative and was taxonomically assigned with the Ribosomal Database Project database (<https://rdp.cme.msu.edu/>) and the classifier program (Wang *et al.* 2007). We assigned each OTU a taxon when the assignment probability was > 80%.

For the whole dataset, we then classified OTUs as following. 1) Any OTUs that did not assign with a probability of 1 to Bacteria could correspond to either degraded or chimeric

fragments generated during the amplification and sequencing process, and was thus defined as artifactual OTUs. Amongst the remaining OTUs, we further identified 2) contaminant OTUs, i.e. having a source other than bacteria found on the sampled ant cuticle, either from chemical reagents and technical material. These were defined as having a maximal abundance in negative controls. The remaining OTUs were defined as 3) the targeted bacterial community. In each biological sample, we further identified 4) false positive OTUs that are generated during the library preparation or the sequencing ((Esling *et al.* 2015; Schnell *et al.* 2015). These were defined as any OTU having a relative abundance in one sample representing less than 0.01% of the OTU total abundance across the entire dataset. Corresponding relative abundances were set to 0 in downstream analyses.

7) *Statistical analyses*

To evaluate potential errors in DNA extraction protocols we assessed whether the extraction methods yielded similar amounts of artifactual, contaminants bacterial or false positives OTUs. To this end, the proportions of artifactual OTUs and contaminants in each sample were log transformed to meet assumptions of normality. We then performed a linear regression using the extraction method as explanatory variable and ant species as a cofactor. Post-hoc pairwise comparisons between extraction methods were performed by removing the ant species effect and conducted with multiple comparisons of means (Tukey contrasts) with FDR-corrected p-values. This analysis was performed by considering all PCR duplicates separately as contamination may occur during the amplification process.

Extraction success was assessed to establish the performance of each DNA extraction protocol. We here summed duplicate PCRs for each individual to determine variation across biological replicates. The number of reads can be indicative of the initial number of 16S copies (Rubin *et al.* 2014) and hence of DNA extraction success, but also of PCR and sequencing success. These were compared, together with the number of OTUs, across DNA extraction protocols, ant species, individual vs. pool, and body size. Because these variables were overdispersed, we used the Generalized Linear Model (GLM) with a negative binomial distribution. The effect of different factors was assessed using a two-way ANOVA. First, we tested the effect of DNA extraction protocols, ant species, and their interaction from samples of single individuals. Next, we tested the effect of number of individuals (single individuals vs. pools of 10 individuals), ant species, and their interaction using the QIAamp method only. Finally, we tested the number of individuals, body size, and their interaction on *A. cephalotes* samples. Post-hoc pairwise comparisons were then performed with Tukey post-hoc tests with FDR p-value corrections for multiple testing.

We further assessed whether the cuticular communities were appropriately sampled by each DNA extraction protocols using rarefaction curves. We conducted this analysis on averaged communities for each method, ant species, as well as for blank controls. Indeed, blank controls, as well as samples for which extraction, PCR or sequencing did fail may still contain several false positive or contaminants OTUs that were not successfully removed during the data curation process. These should result in flatter rarefaction curves and lower number of OTUs than for biological samples.

We then assessed the effect of DNA extraction protocols, number of individuals used for extraction, and individual body size on the cuticular bacterial beta diversity (community composition variation across samples). To this end, we excluded samples having less than 100 reads (n=12) as they most likely correspond to failed PCRs. We did not rarefy the data in our analysis because we considered the number of reads produced to be part of the evaluation criteria. Beta diversity patterns were obtained by computing pairwise Bray-Curtis dissimilarities on Hellinger-transformed data (Legendre & Legendre 2012) using non-averaged samples. Community differences were visualized using a Principal Coordinates Analysis (PCoA), and effects of DNA extraction method, ant species, and their interaction were tested with a Permutational Multivariate Analysis of Variance (PERMANOVA) (Anderson 2001).

The multivariate homogeneity analysis of group dispersions (Anderson 2006) for each DNA extraction protocol and ant species was performed to assess the reproducibility of the retrieved bacterial communities composition in each method. Differences in reproducibility were tested with Tukey post-hoc tests with FDR p-value corrections.

We further examined whether extraction methods were biased toward particular bacterial clades. To this end, we computed the bacterial clades log2 fold-change ratio between DNA extraction methods (Rapaport *et al.* 2013). We added one read to all OTUs abundances prior the analysis to avoid null values. Significance of change was assessed with a Mann-Whitney test and p-values were corrected with the FDR method.

Finally, we performed similar analyses of beta diversity described above in order to test (i) the effect of pooling, ant species and their interaction and (ii) the effect of pooling and body size.

All analyses were carried out in R (<https://www.R-project.org/>) with R packages *vegan* (<https://cran.r-project.org/web/packages/vegan/index.html>), *MASS* (<https://cran.r-project.org/web/packages/MASS/index.html>), and *RVAideMemoire* (<https://cran.r-project.org/web/packages/RVAideMemoire/index.html>).

D. Results

1) Comparison of DNA extraction protocols

Overall, low DNA concentrations were retrieved in the washing solutions of ant cuticles. For most *A. cephalotes* samples DNA contents could be measured with the Qubit fluorometer, however it was not possible to discriminate between extraction methods (data not shown). The DNA quantities extracted from all *P. penetrator* samples were below the detection threshold of the Qubit fluorometer (0.02 ng/μL).

(a) Artifactual sequences and DNA contaminants

We identified a total of 2,277 artifactual OTUs. They represented 18.8% of the total number of OTUs and 1.9% of reads in the dataset. The proportion of artifactual OTUs in samples differed between DNA extraction methods (ANOVA $F_{3,191} = 9.789$, $p < 0.001$) regardless of ant species (ANOVA $F_{1,191} = 0.916$, $p = 0.34$; species $\times$ extraction: ANOVA $F_{3,191} = 1.196$, $p = 0.313$). Dneasy yielded consistently less artifactual sequences compared to the other methods for both ant species according to post-hoc pairwise comparisons (Table S1).

We also retrieved a total of 142 contaminant OTUs (Table S2). They corresponded to 1.2% of all OTUs and 0.2% of the total number of reads. Only five of these had a maximum abundance in PCR controls and represent less than 200 reads among all samples. Their rarity and lack of consistent amplification across controls suggest sporadic air contaminations rather than contamination present in PCR reagents. Seven contaminant OTUs had a maximal abundance in tweezer controls and represented more than 1000 reads each. They were affiliated to different bacterial taxa; the most abundant one being assigned to the genus *Alkalibacterium*.

Excluding these 12 contaminant OTUs, we retrieved 130 more contaminant OTUs of which maximum abundance was found in DNA extraction controls. Amongst the most recurrent bacterial contaminant OTUs (i.e. occurring in 25% of all extraction controls), we identified members of Gammaproteobacteria belonging to Xanthomonadaceae and Halomonadaceae. When the contribution of contaminant OTUs to sample diversity across *A. cephalotes* and *P. penetrator* individuals was examined we found that both ant species (ANOVA $F_{1,191} = 26.539$, $p < 0.001$), extraction methods (ANOVA $F_{3,191} = 16.305$, $p < 0.001$) and their interactions (ANOVA $F_{3,191} = 4.252$, $p = 0.006$) had an effect on these proportions, with a particularly high amount of contaminant OTUs in *P. penetrator* samples (Table S1). For these samples, Dneasy and MO BIO yielded consistently less contaminant OTUs according to post-hoc pairwise comparisons (Table S1).

We also looked for the number of false positives OTUs in biological samples. The average proportion of reads obtained with unused tag combinations compared to biological samples corresponded to 0.21%, hence suggesting a relatively low false positive rate in our data. Nevertheless, we examined the proportion of reads corresponding to false positive OTUs in *A. cephalotes* and *P. penetrator* samples. We found that both ant species (ANOVA $F_{1,191} = 17.565$, $p < 0.001$) and extraction methods (ANOVA $F_{3,191} = 13.964$, $p < 0.001$) had an effect (interaction: ANOVA $F_{3,191} = 2.504$, $p = 0.06$), with a higher proportion of reads corresponding to false positive OTUs in *P. penetrator* samples, and in samples extracted with Dneasy for both ant species (Table S1).

(b) Effect on the number of reads, OTUs and diversity coverage

After removing contaminants and artifactual OTUs, the dataset comprised 9.687 OTUs and 3,553,385 reads. When comparing the numbers of reads and OTUs across samples, both ant species (ANOVAs, $X^2 = 76.713$ for reads and $X^2 = 156.895$ for OTUs, $df = 1$, $p < 0.001$ in both cases) and extraction methods had significant effects (ANOVAs, $X^2 = 22.474$ for reads and $X^2 = 24.535$ for OTUs, $df = 3$, $p < 0.001$ in both cases) but not their interaction (ANOVAs, $X^2 = 6.427$ for reads and $X^2 = 7.548$ for OTUs, $df = 3$, $p = 0.1$ in both cases). *A. cephalotes* samples contained more reads and OTUs than *P. penetrator* samples for all extraction methods (Fig. 2, Table S3, Tukey post-hoc test, most $p < 0.001$). Dneasy yielded significantly less reads and OTUs for *P. penetrator* (Table S3, Tukey post-hoc test, all $p < 0.001$), and even produced less reads/OTUs than some extraction controls (Figure 40). PC yielded significantly less OTUs for both species than QIAamp (Table S3, Tukey post-hoc test, all $p = 0.01$). PC also yielded a similar number of reads and OTUs for *P. penetrator* samples and extraction controls. MO BIO produced higher number of reads and OTUs in *P. penetrator* samples compared to controls. Performing the analysis on non-curated data did not change the results (Fig. S1).

Rarefaction curves were drawn on averaged communities per extraction method and species in order to determine whether the cuticular bacterial diversity was sufficiently sampled. Although none of the rarefaction curves did reach a plateau for biological samples (Figure 40), we found noteworthy differences. First Dneasy produced the shortest rarefaction curves for both ant species samples while PC displayed the flattest curves and the lowest number of OTUs for *A. cephalotes*. For both ant species, the samples extracted with QIAamp and MO BIO yielded rarefaction curves that were clearly different from blank control curves and yielded more OTUs.

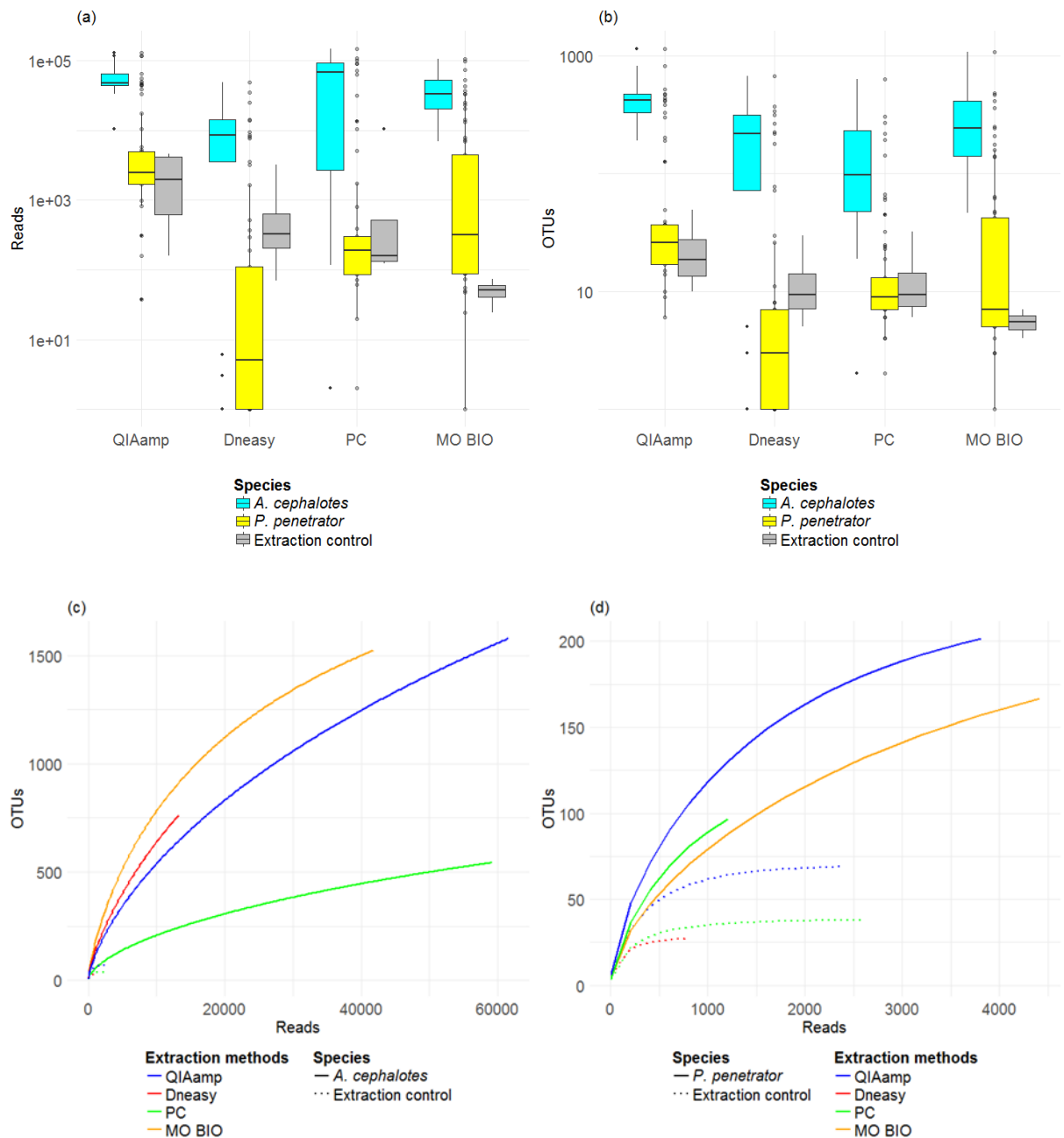


Figure 40 : Effect of DNA extraction protocols on the number of reads (a), OTUs (b) and diversity coverage (c-d). In (a-b), number of reads and OTUs are represented for *A. cephalotes* (blue), *P. penetrator* (yellow) ant species with cleaned data and a logarithmic scale (y axis). The figure includes DNA extraction controls (grey) for comparative purposes. (c-d) represent the rarefaction curves on averaged bacterial communities for each DNA extraction methods for (a) *A. cephalotes* and (b) *P. penetrator*.

(c) Taxonomic composition, and beta diversity

The most abundant phyla found on the studied ants were Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Firmicutes, Bacteroidetes and Actinobacteria, and the abundances of these phyla tend to differ between ant species (Figure 41 a) with

Actinobacteria being more abundant in *A. cephalotes* and Betaproteobacteria and Bacteroidetes being more abundant in *P. penetrator*.

These differences were confirmed at the OTU level as *A. cephalotes* and *P. penetrator* bacterial communities clearly differed along the PCoA axis 1 in (Figure 41 b). Differences between DNA extraction methods were also observed along the PCoA axis 2 for samples of both species and supported by the PERMANOVA analysis (species: $F_{1,80}=9.046$, $p=0.001$, $R^2=0.091$; extraction method: $F_{3,80}=3.563$, $p=0.001$, $R^2=0.107$, interaction: $F_{3,80}=2.297$, $p=0.001$, $R^2=0.069$). The effect of DNA extraction methods was consistent when repeating the analysis in each ant species separately (*A. cephalotes*: $F_{3,46}=2.673$, $p=0.001$, $R^2=0.157$ *P. penetrator*: $F_{3,33}=3.192$, $p=0.001$, $R^2=0.242$). *P. penetrator* samples were divided into two groups: a first group of samples extracted with PC and MO BIO and a second group of samples extracted with QIAamp and Dneasy (Figure 41 b). The pattern is less clear in *A. cephalotes* samples, where the DNA extraction methods are loosely clustered and separated into different groups.

The repeatability of each extraction method was similar amongst methods for both ant species, as we did not find significant differences in the dispersion of community composition (Fig. S2, Tukey post-hoc test, all $p>0.1$). However, the distances to centroids tended to be more variable for PC and MO BIO for both ant species. When repeating the analysis in each ant species separately, QIAamp displayed better repeatability than PC and MO BIO for *P. penetrator* samples (Tukey post-hoc test, $p=0.01$).

To further characterize potential biases in the bacterial community composition analysis from DNA extraction protocols, we assessed the differential abundance of the most abundant bacterial phyla between methods (Fig. S3). Overall, we did not find strong and systematic taxonomic biases within the methods tested.

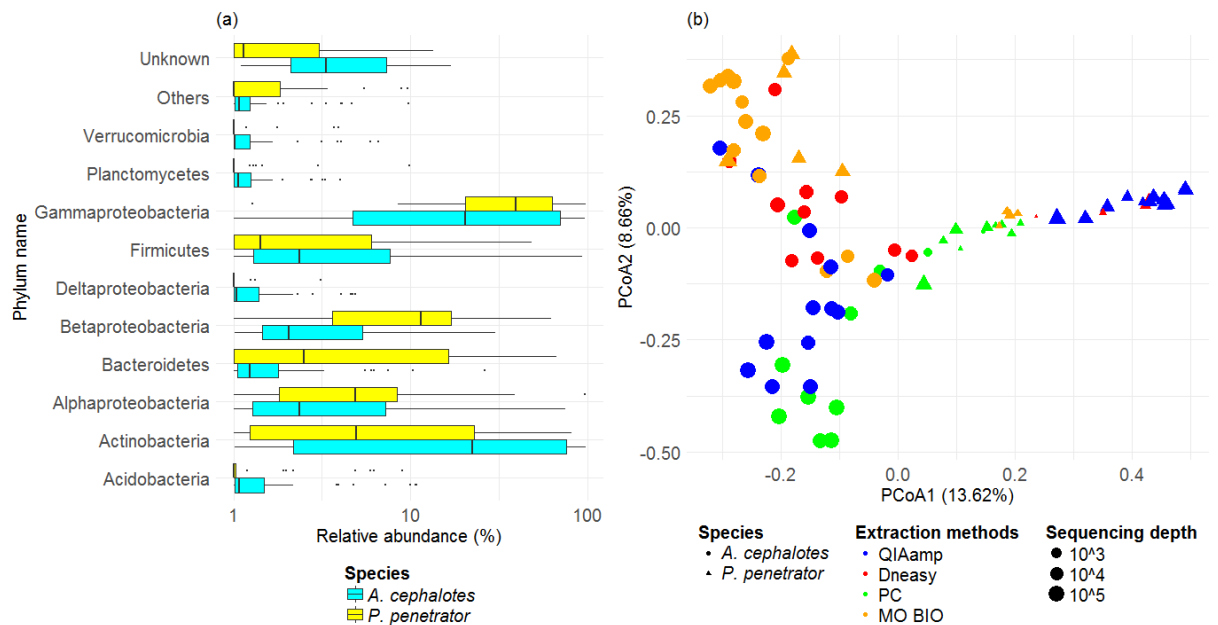


Figure 41 : Taxonomic composition of bacterial cuticular microbiota and effect of DNA extraction methods on bacterial beta diversity. (a) Taxonomic composition of bacterial cuticular microbiota of *A. cephalotes* (blue) and *P. penetrator* (yellow). Only bacterial clades with a total relative abundance greater than 0.1% are shown. (b) Similarities of samples in their community composition according to DNA extraction methods. Dot size represents the sequencing depth of each sample. Only samples with more than 100 reads are included.

2) Effect of pooling individuals and body size on cuticular diversity

We assessed the effect of number of individuals and individual body size on the cuticular microbiota by using QIAamp only, as it both yielded higher number of reads and one of the most consistent signals compared to other methods. DNA concentrations were the highest in pools of individuals compared to individual alone regardless the ant species (Fig. S4a). For *A. cephalotes* samples, pools of majors also yielded more DNA than pools of minors (Fig. S4d). Pools of individuals yielded significantly more OTUs than single individuals (ANOVA, $df = 1$, $X^2 = 29.716$, $p < 0.001$) in both *A. cephalotes* (1460 ± 659 vs. 521 ± 270 OTUs respectively) and *P. penetrator* (143 ± 84 vs. 35 ± 32). However, the number of reads was not affected by the number of individuals (ANOVA, $df = 1$, $X^2 = 1.647$, $p = 0.20$; Fig. S4b,c). *A. cephalotes* pools of minors had less OTUs (175 ± 91) than majors alone or pools of majors (Tukey post-hoc test, all $p < 0.001$), despite yielding similar amounts of reads (number of individuals: ANOVA, $df = 1$, $X^2 = 0.298$, $p = 0.58$ and size: ANOVA, $df = 1$, $X^2 = 0.928$, $p = 0.33$; Fig. S4e,f). The rarefaction curves of averaged communities obtained from single individual vs. pooled of 10 individuals were similar for both species (Fig. S5a). On the other hand, bacterial diversity of pools of *A. cephalotes* minors was much lower, and hence better covered compared to that of majors (Fig. S5b).

We finally explored how these factors may influence bacterial community composition. The PCoA ordination and PERMANOVA analysis showed net differences between *A. cephalotes* and *P. penetrator* along the PCoA axis 1, as well as depending on the number of individuals extracted (Figure 42 a; PERMANOVA species: $F_{1,32}=11.482$, $p=0.001$, $R^2=0.246$; number of individuals: $F_{1,32}=3.07$, $p=0.002$, $R^2=0.066$, interaction: $F_{1,32}=3.01$, $p=0.002$, $R^2=0.065$). The effect of the number of individuals extracted was consistent when repeating the analysis for each ant species separately (PERMANOVAs *A. cephalotes*: $F_{1,16}=2.544$, $p=0.002$, $R^2=0.145$; *P. penetrator*: $F_{1,15}=3.823$, $p=0.002$, $R^2=0.214$). In addition, *A. cephalotes* samples formed two groups according to the body size of the extracted individuals along the PCoA axis 2 (Figure 42 a, PERMANOVA number of individuals: $F_{1,20}=3.557$, $p=0.001$, $R^2=0.135$; size: $F_{1,20}=4.704$, $p=0.001$, $R^2=0.179$). These differences held when taking into account differences in sequencing depth amongst pools of individuals differing in their body size (PERMANOVA size: $F_{1,7}=7.566$, $p=0.01$, $R^2=0.514$; sequencing depth: $F_{1,7}=2.159$, $p=0.149$, $R^2=0.147$, interaction: $F_{1,7}=0.981$, $p=0.364$, $R^2=0.067$).

The most abundant phyla found in pools of *A. cephalotes* minors and majors were Gammaproteobacteria, Actinobacteria, Alphaproteobacteria, Betaproteobacteria, Firmicutes, Bacteroidetes, but the relative abundances of these phyla tend to differ between the two size classes (Figure 42 b).

We finally assessed again the repeatability of the retrieved bacterial community for individuals and pools of individuals, as well as depending on individual body size. Bacterial communities were more variable in *A. cephalotes* majors when individuals were extracted separately (Tukey post-hoc test, $p=0.01$, Figure S6). We did not observe such differences in *P. penetrator* (Tukey post-hoc test, $p=0.62$).

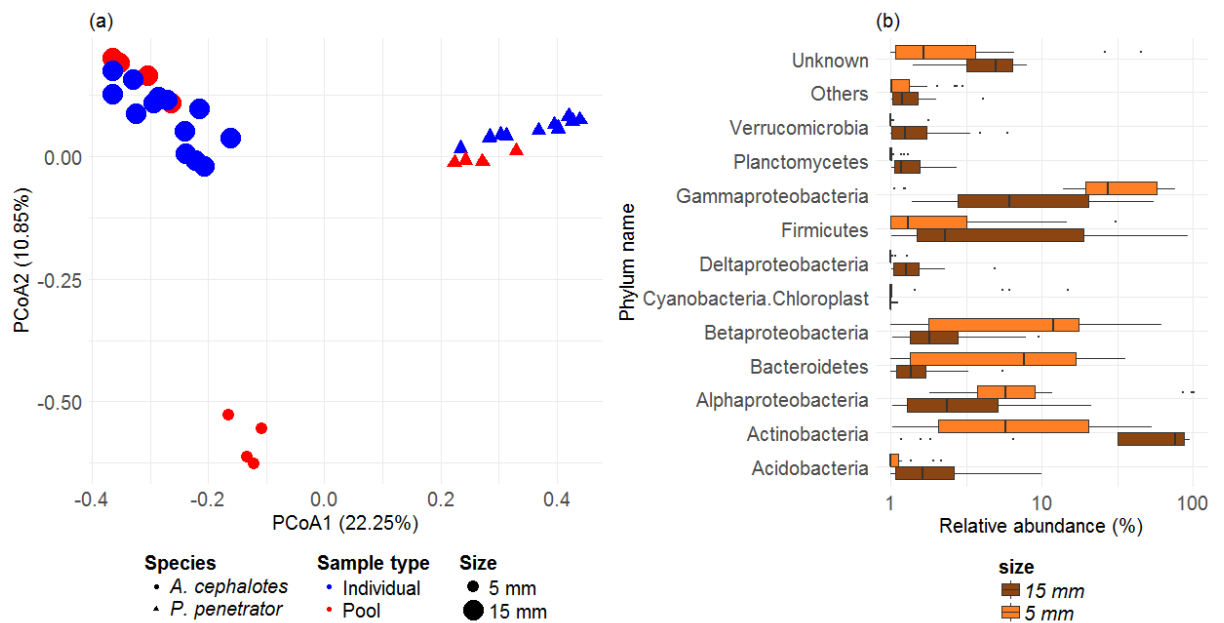


Figure 42 : Effect of the number of individuals and individual body size on bacterial beta diversity and taxonomic composition of bacterial cuticular microbiota. (a) Bacterial community composition similarities between *A. cephalotes* and *P. penetrator* using QIAamp DNA extraction protocol. (b) Taxonomic composition of bacterial cuticular microbiota of *A. cephalotes* major workers (brown) and minor workers (orange). Only bacterial clades with a total relative abundance greater than 0.1% are shown.

E. Discussion

DNA extraction protocols have been extensively benchmarked on widely diverse matrices including soil (Zinger *et al.* 2016), permafrost (Thomsen & Willerslev 2015), faeces (Wesolowska-Andersen *et al.* 2014), and insect guts (Rubin *et al.* 2014). While such samples all have particularities that require optimizations, they contain non-limiting bacterial biomass, which makes bacterial DNA extraction relatively straightforward. This is, however, not the case for the low bacterial loads retrieved in washes of body surface insects, which are more prone to contaminations and to yield artifactual sequences after PCR and sequencing. Here, we evaluated four DNA extraction methods in their ability to maximise the strength and robustness of the estimated bacterial community diversity retrieved from the cuticle of two ant species. We also assessed whether the ant size, or the number of individuals would affect the diversity and composition of the retrieved bacterial communities.

1) Effect of DNA extraction protocols on the cuticular bacterial diversity

We studied the cuticular microbiota of two ant species that strongly contrast in their body size and ecology. They may therefore display variable degrees of association with bacteria

or bacterial loads. Accordingly, the two-species differed in the yield of DNA retrieved after extraction, extraction/PCR/sequencing success, as well as in the diversity and community composition of their cuticular bacterial communities. More specifically, the largest and ground-nesting ant species (*A. cephalotes*) systematically yielded larger amounts of bacterial DNA and richer bacterial communities compared to the arboreal obligate myrmecophyte one (*P. penetrator*), irrespective of the DNA extraction protocols. While the causes of these differences (i.e. a sampling effect due to body size differences, exposure to environmental sources, or actual differences in the strength of their association with bacteria) are beyond the scope of this paper, the differential amplification of bacteria illustrates that our experimental design included samples that differed in their difficulty of success: extracted DNA from *A. cephalotes* was often quantifiable, while extracted DNA from *P. penetrator* samples was systematically under the Qubit detection threshold. We found that the *A. cephalotes* DNA extracts contained less contaminants and false positives than those of *P. penetrator*. Such finding agrees with the expectation that the degree of contamination and amount of false positive in samples is inversely correlated with the initial DNA content, as reported in studies on microbial communities of paediatric bronchoalveolar lavages (Willner *et al.* 2012) and on false positive rates (Esling *et al.* 2015).

This emphasizes the importance of using negative controls when working with such sensitive samples. Indeed, we retrieved systematically DNA reads in negative controls, in agreement with previous reports (Ishak *et al.* 2011; Salter *et al.* 2014), which are mainly composed of remaining contaminants or to false positive OTUs. These enable us not only to identify many contaminants and remove them in subsequent analyses, but also to ensure that bacterial communities of *P. penetrator* were not artifactual, i.e. presented a minimal overlap with negative controls, as done in a previous work on the microbiota of the ant *Solenopsis invicta* (Ishak *et al.* 2011).

After data cleaning, we observed noticeable differences amongst the four DNA extraction methods tested in terms of extracted DNA yield, number of sequencing reads retrieved, and bacterial community diversity and composition that were consistent between the two species in most cases, and particularly more visible in *P. penetrator* samples. Such differences have been commonly reported elsewhere e.g. for host-associated microbiota (Willner *et al.* 2012; Rubin *et al.* 2014), soils (Zinger *et al.* 2016), or faecal samples (Wesolowska-Andersen *et al.* 2014) and may arise from methodological inconsistencies in either the efficiency of bacterial cells lysis or the efficiency in eluting DNA in downstream analyses.

Amongst the methods benchmarked, Dneasy yielded the lowest amounts of sequencing reads, OTUs as well as the highest amounts of false positives. This suggests that Dneasy was the least efficient for extracting ant cuticle microbiota. The truncated form of the Dneasy rarefaction curve for *A. cephalotes* species further supports this conclusion. Although we cannot exclude that the truncated form is due to remaining false positive OTUs that we were not able to filter out, the curves were strongly distinct from those obtained with extraction controls. On the other hand, the rarefaction curve obtained for *P. penetrator* could not be distinguished from controls samples. Hence, we suspect Dneasy to grossly under-samples bacterial diversity, and that its inefficiency may lie in a significant DNA loss during DNA capture or elution on the silica-membrane, which has larger pore size in Dneasy kit compare to QIAamp.

The PC method also yielded significantly lower numbers of OTUs in both ant species compared to QIAamp. PC rarefaction curves also indicated that bacterial diversity was relatively well sampled in both ant species through sequencing. We hence suspect that the PC method also under-sampled bacterial diversity. This contrasts with the idea that PC yields more bacterial DNA (Rubin *et al.* 2014). We explain this difference either by an incomplete cell lysis for some bacteria, or by an inhibition of the PCR by phenols, which may have greater effects when working with low amounts of DNA.

On the other hand, QIAamp and MO BIO were more efficient. They yielded higher amounts of reads and OTUs for both ant species, similar rarefaction curves that clearly differed from extraction blank controls. However, bacterial communities retrieved with MO BIO were more variable for *P. penetrator*. In addition, they produced different bacterial communities, with those obtained with MO BIO being similar to those produced with PC. Although we cannot exclude that MO BIO yields a realistic biological signal for *P. penetrator*, QIAamp displayed higher consistency and repeatability, produced higher levels of contamination and artifactual OTUs. Our result hence suggest that this method is a good compromise to retrieve the cuticular microbiota on single individuals of various insect species, even in the small-bodied *P. penetrator*, provided that negative controls and appropriate bioinformatic treatments are included in the experiment to detect and remove artifactual and contaminant OTUs.

2) *Impact of number of individuals and individual body size*

To insure retrieving sufficient bacterial DNA, previous studies recommended to conduct DNA extraction methods on pools of ant individuals (Seipke *et al.* 2013). However, the inter-individual variability of cuticular microbiota can be of importance. For example, fishes with

certain genotypes have been found to be more associated with particular beneficial skin bacteria (Boutin *et al.* 2014) Furthermore, replication of bacterial community signals among individuals reinforces the notion of regulated environments rather than communities composed of random opportunists. We, therefore evaluated whether bacterial communities retrieved from single individuals compared with those retrieved from pools of individuals. We found that extraction of single individuals produced lower DNA yields and OTUs than that of pools of 10 individuals for both ant species, hence suggesting that indeed, pooling individuals would maximize the amount of bacterial DNA extracted. However, we also found that the diversity and bacterial community composition was to less well covered and much more variable on single individual samples. Since the number of reads retrieved per sample did not differ between individual and pooled samples, these results strongly suggest that we were able to capture the intraspecific individual variability in cuticular microbiota in both ant species. This property would hence allow better understanding the mechanisms shaping insect's cuticular microbiota, which is particularly relevant in social insects as it may determine the health and fitness of the whole colony.

Insect body size can be viewed as a proxy of insect cuticle surface area, and is obviously an important limitation that challenges the sensitivity of DNA extraction methods. As such, cuticular bacterial DNA quantities are positively correlated with insect body weight (Ushio *et al.* 2015). Our results are in agreement with this trend, with major *A. cephalotes* having more DNA extracted and OTU numbers than minor *A. cephalotes* and *P. penetrator* workers (5 mm). This supports our previous finding and again emphasizes the need to adopt rigorous experimental practices when working with small insects to minimize contaminants and PCR/sequencing artefacts at both the molecular and bioinformatic steps.

Beyond the effects of insect body size on the amount of DNA retrieved, our study further shows that insects of different body size within a single species, and even within a single colony can impact bacterial community composition to the same extent as if two different insect species were compared. In leaf-cutting ants, the social structure of the colony is composed of four physical castes. The body size and age are related to the distribution of tasks among workers within the colony. Typically outside tasks are performed mainly by older individuals (Wilson 1980). The minor and major ant individuals studied here were collected outside the nest and belonged to two different castes : 1) the minors riding atop the leaves that guard against parasitic flies attempting to attack and lay eggs in the leaf-carrying media's necks; and 2) the majors, who are defenders and constitute a distinct physical caste, called "soldiers" (Wilson 1980).

We observed major differences between these two castes, suggesting a determinist mechanism in the shaping of bacterial communities associated with *A. cephalotes* cuticles. In particular, we found that majors harboured more Acidobacteria, Verrucomicrobia and Planctomycetes. These are commonly of soil and litter environments (Fierer *et al.* 2009) and may reflect majors' contact to the forest floor. On the other hand, Alphaproteobacteria, Bacteroidetes, and Betaproteobacteria were more abundant in the cuticular microbiota of Atta minors. These groups are important components of the phyllosphere microbiota (Kembel *et al.* 2014), and minors may have acquired them through their contact with leaves.

We also noticed a greater relative abundance of Actinobacteria in Atta majors samples. The ant tribe Attini are engaged in a mutualistic interaction with antimicrobial metabolites producing Actinobacteria on their cuticle (Currie *et al.* 1999b). Recent work has demonstrated that newly emerged workers of the *Acromyrmex* genus are symbiont-free and must acquire cuticular Actinobacteria communities from major workers (Marsh *et al.* 2014). The differences observed here might hence result from such mechanism. Of course, these hypotheses should be confirmed with further experiments.

In conclusion, our benchmark suggests that despite the methodological challenges, the use of QIAamp together with a rigorous controls and data curation processes allow retrieving reliably the inter-individual variability of the cuticular microbiota, which might be dramatically increased by insects body size or behaviour differences. The capacity of evaluating individual cuticular microbiomes hence opens new perspectives in determining the eco-evolutionary processes involved in insect cuticle-microbiome interactions.

F. Annexes : Figures et Tables du SI

L'ensemble des appendix, figures et table du *Supporting information* de l'article sont en Annexe 6.

Chapitre 3. Jardin de fourmis : Entre microbiote et métabolome cuticulaire.

I. Introduction

Le rôle majeur attribué au microbiote cuticulaire des insectes est un rôle de défense contre les infections engendrées par des entomopathogènes. Bien qu'il soit admis que le microbiote cuticulaire contient une plus forte proportion de micro-organismes bénéfiques que par chance (Scheuring & Yu 2012), des zones d'ombres subsistent quant à l'assemblage du microbiote cuticulaire bénéfique des insectes sociaux: est-il le fruit d'un processus co-évolutif et donc transmis verticalement ou est-il acquis à partir de l'environnement bactérien dans lequel les insectes vivent ? L'étude de Scheuring & Yu suggère que ces deux processus entrent en jeux. D'après ce travail, une explication plausible de l'association à long terme entre les attines et leurs symbiontes microbiens spécifiques et sans apparition de résistance chez les pathogènes serait que seulement quelques souches sont transmises verticalement et favorisent le recrutement de plusieurs taxons microbiens environnementaux. La compétition entre ces derniers taxons stimulerait la synthèse de cocktails d'antimicrobiens très diversifiés.

Cette dernière idée est supportée par nos résultats du premier chapitre, où nous avons montré que certains métabolites antimicrobiens exprimés par les actinobactéries cultivées en coculture sont plus abondants et plus divers qu'en monoculture du fait des interactions microbiennes (c'est à dire compétition, commensalisme, communication). Il est même probable que nous sous-estimions la diversité de ces cocktails *in natura*, puisque notre analyse chimique des cocultures ne représente que la part du métabolome des bactéries cuticulaires cultivables. D'autre part, le chapitre 2 nous a permis de montrer des différences significatives dans des communautés bactériennes de la cuticule de deux espèces de fourmis aux traits éco-physiologiques distincts. Ces 2 espèces ne partageant pas le même habitat, il ne nous est pas possible de savoir si les communautés microbiennes sont spécifiques à l'espèce ou à leur habitat. Néanmoins, nous avons également pu mettre en évidence de la variabilité interindividuelle du microbiote cuticulaire au sein d'une colonie de la même espèce en particulier entre deux différentes castes chez l'espèce *Atta cephalotes*. Ceci suggérerait que le microbiote cuticulaire est acquis horizontalement, mais l'échantillonnage de cette étude étant limité à une seule colonie, cette expérience ne nous permet pas de conclure. Afin d'apporter de nouveaux éléments à la compréhension des mécanismes de transmission des communautés bactériennes cuticulaires chez les fourmis et de tenter d'accéder aux métabolites du microbiote bactérien cuticulaire *in situ* nous avons mené dans ce troisième chapitre, l'étude du microbiote

en parallèle de celle du métabolome cuticulaire de deux espèces de fourmis collectées dans plusieurs sites de prélèvement à une échelle régionale.

Plus précisément, nous proposons ici d'étendre notre échantillonnage en collectant en parallèle deux espèces de fourmis. Nous avons choisi comme modèle d'étude les jardins de fourmis constitués de deux espèces de fourmis, *Ca. femoratus* et *Cr. levior*, qui partagent le même nid localisé dans le réseau racinaire de plusieurs épiphytes. La variabilité de la composition des communautés bactériennes liée à des différences nettes d'habitats entre espèces est donc ici contrôlée puisque ces deux espèces de fourmis partagent le même nid. De plus, ce couple d'espèce se retrouve à des endroits géographiques variés et associé à différents groupes d'épiphytes, en d'autres termes dans des environnements microbiens variés. Cet exemple étonnant de parabiose est donc un cas d'étude de choix car il permet de tester si le microbiome cuticulaire résulte de la transmission verticale ou de la variabilité du pool bactérien environnemental. Par exemple, s'il existe des différences de compositions bactériennes entre les sites indépendamment de l'espèce, ceci renforce l'hypothèse de l'acquisition des souches dans l'environnement. Dans le cas contraire, ceci renforcerait l'hypothèse de la transmission verticale. Aux vues des travaux de Scheuring & Yu, nous nous attendons à ce que la majorité du microbiote cuticulaire soit acquis horizontalement. Les symbiontes spécifiques à l'espèce, et donc vraisemblablement transmis verticalement, devraient constituer une minorité du microbiote. Dans le cadre de la parabiose étudiée, nous déterminerons également s'il existe des taxons microbiens partagés par les deux espèces indépendamment du site de prélèvement, car ils pourraient être impliqués dans le maintien de la parabiose (e.g. communication, tolérance, protection jointe).

Les hypothèses concernant le métabolome cuticulaire de ces deux espèces sont moins évidentes. Néanmoins, sachant que le métabolome cuticulaire des fourmis est un produit du microbiote et des glandes cuticulaires de la fourmi, notre hypothèse est que toute modification de la composition du microbiote entraînerait des modifications de la composition du métabolome. Par exemple dans le cas d'une acquisition du microbiote horizontale, i.e. du pool d'espèces bactériennes environnantes, la composition en métabolite devrait différer entre les sites de prélèvement également. Nous cherchons donc à identifier quels sont les facteurs influant sur les modifications de la composition du microbiote et si ces mêmes facteurs impactent la composition du métabolome. Pour identifier la part des métabolites microbiens bioactifs sur la cuticule il sera nécessaire (i) d'accéder à l'intégralité des métabolites présents *in situ* et (ii) d'utiliser les outils d'identifications structurales (MS/MS) et leurs bases de données (origine bactérienne et activité antimicrobienne documentée) pour discriminer les métabolites

d'origine microbienne au sein du métabolome cuticulaire. Cette approche permettrait de préciser la covariation entre le microbiote et ses produits métaboliques, et à terme de découvrir de nouvelles molécules bioactives.

Notre approche permettrait donc de décrire à la fois l'ensemble des métabolites cuticulaires détectables par UPLC-HRMS et l'ensemble des bactéries, cultivables et non cultivables, par séquençage à haut débit. A notre connaissance c'est la première fois que cette méthodologie est testée dans le domaine de la biologie des insectes, contrairement aux études centrées exclusivement sur les métabolites biologiquement actifs isolés à partir de bactéries cultivables (Currie *et al.* 1999b; Gao *et al.* 2014). L'objectif final de ce travail est de traiter de manière conjointe et de relier certains métabolites à la présence de certains OTUs. Dans le cas où les taxons bactériens seraient connus pour la production de métabolites antibactériens ou que les métabolites seraient connus ou préalablement décrits dans le monde microbien, ces résultats permettraient de prouver la fonction de défense du microbiote cuticulaire bénéfique. Ce travail n'a pas pu être totalement réalisé dans le cadre de ce manuscrit pour des contraintes de temps et seule l'analyse en parallèle des données des communautés microbienne et du métabolome est présentée.

II. Partie expérimentale

A. Prélèvements et préparation des échantillons

Trois sites ont été choisis en Guyane pour cette étude, où deux jardins sont prélevés sur chaque site (six jardins au total) : la montagne de Kaw, le barrage de Petit Saut et la station scientifique des Nouragues (Figure 43). Dans chaque jardin de fourmis, dix individus de chaque espèce (*Ca. femoratus* et *Cr. levior*) ont été collectés afin de caractériser leur microbiote par métabarcoding, et leur métabolome cuticulaire par métabolomique. Dans chaque jardin de fourmi, nous avons aussi prélevé quatre échantillons d'un mélange de terre et de racine constituant le « nid » de ces fourmis.

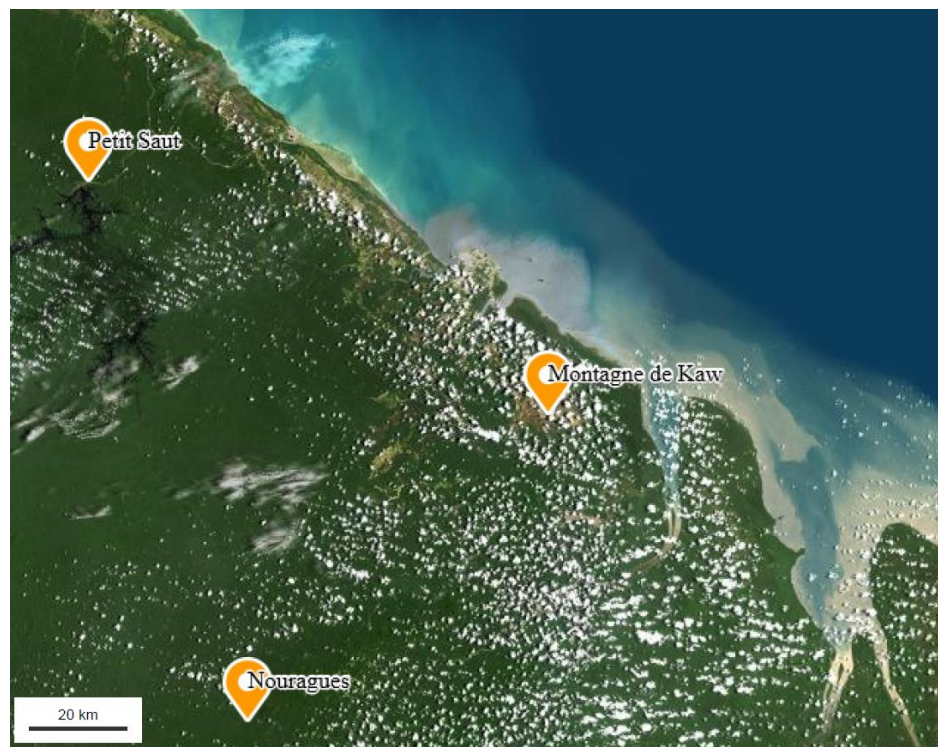


Figure 43 : Carte de Guyane présentant les trois sites de collecte, <https://www.geoportail.gouv.fr/mentions-legales>.

Au total, l'analyse métabolomique a été menée sur 60 individus de *Ca. femoratus* et sur 60 individus de *Cr. levior*. L'analyse métabarcoding a été faite sur d'autres individus ; 60 de *Ca. femoratus*, 60 de *Cr. levior* ainsi que sur 24 échantillons de nid. La rapidité de traitement des prélèvements est essentielle pour les deux études pour éviter toutes modifications du microbiote et du métabolome liées à l'activité microbienne qui pourrait avoir lieu pendant la conservation des échantillons. Le protocole de collecte et de lavage des individus dans le solvant et la solution tampon se sont donc déroulés le même jour et effectués sur le terrain si le site était trop éloigné du laboratoire de Cayenne.

Pour l'analyse métabolomique, les fourmis ont été introduites individuellement dans des tubes en plastique contenant un mélange de solvant. Ce mélange de solvant acétate d'éthyle/méthanol/eau (2:1:1) permet de solubiliser un maximum de métabolites aux polarités différentes, des plus hydrophiles aux plus lipophiles. Les échantillons ont été agités vivement (vortex, 1 min) pour faciliter la solubilisation des métabolites de la cuticule dans le mélange de solvant. La solution obtenue a été ensuite introduite dans des contenants en verre (*vials*) qui sont ensuite conservés à -80°C jusqu'à l'injection ultérieure dans un spectromètre de masse.

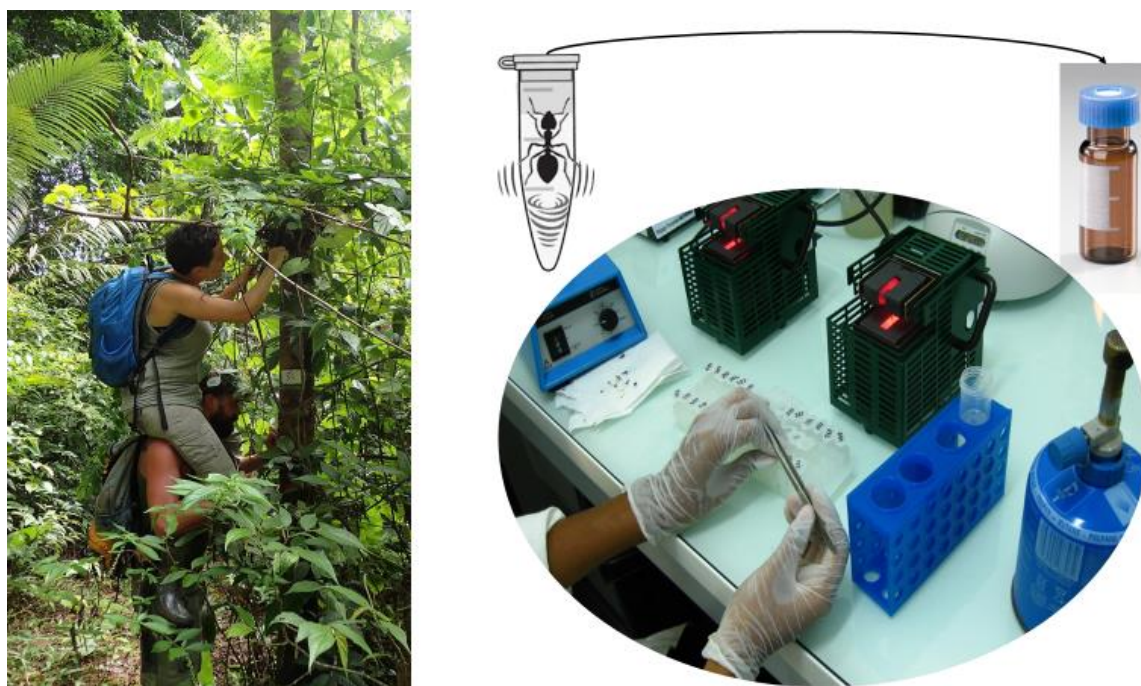


Figure 44 : Photographies illustrant les étapes du prélèvement et du lavage, (auteur : Christophe Duplais).

Pour l'analyse métabarcoding, les fourmis ont été introduites individuellement dans des tubes en plastique contenant un tampon de lyse A. Le tampon de lyse contient un surfactant (Triton X-100) permettant de détacher les micro-organismes de la cuticule. Il contient aussi du lysozyme et de la protéinase K qui permettent de lyser les cellules bactériennes. Les échantillons ont été agités vivement (vortex, 1 min) pour améliorer le détachement des bactéries et la lyse de cellules. Les fourmis ont été manipulées avec précaution à l'aide de pinces stérilisées à la flamme et à l'alcool pour ne pas introduire de contaminants (Figure 44). Les eaux de lavage ont été conservées à -20°C avant l'extraction de l'ADN.

B. Analyse par métabolomique

1) Acquisition des données

L'analyse métabolomique effectuée lors de ce travail de thèse repose sur la chromatographie UPLC couplée à un spectromètre de masse de haute résolution (*High*

Resolution Mass Spectroscopy) (UPLC-HRMS). Dans le cadre des travaux présentés ici, le spectromètre de masse SYNAPT G2 HDMS de la société Waters a été utilisé. Ces analyses ont été conduites en août 2016 par Grégory Genta-Jouve au laboratoire de toxicologie de l'Université Paris Descartes. Brièvement, les 120 échantillons des métabolites cuticulaires de fourmis (60 *Ca. femoratus* et 60 *Cr. levior*) sont injectés aléatoirement en UPLC-HRMS (spectromètre de masse = QTOF). En parallèle, un échantillon de contrôle qualité (*Quality Control*, QC) est injecté tous les 5 échantillons tout le long de l'analyse. Le QC est un pool représentatif de l'ensemble des échantillons qui permet de suivre la stabilité du système tout au long de l'analyse. Un problème technique a empêché l'injection des derniers échantillons. Les données de métabolites cuticulaires ont donc été acquises pour 51 *Ca. femoratus* et 51 *Cr. levior*.

En mode d'ionisation positive, 5 µL de l'échantillon sont injectés sur une colonne (Waters Acquity UPLC, 2,1*100 mm, porosité 1,7 µm) à 40°C. La phase mobile délivrée dans un gradient d'élution de phase inverse à 0,3 mL/min avec de l'eau (éluant A) et de l'acétonitrile (éluant B) les deux contenant 0.1% d'acide formique. Le gradient de solvant suivant a été utilisé : 5% de B à 100% de B (0-7min). Le spectromètre de masse avec une acquisition de 100 à 1500 m/z. Les données sont collectées en mode centroïde.

Des étapes de prétraitement des fichiers bruts de données sont effectuées afin d'extraire les variables d'intérêt du bruit et leurs informations pour l'interprétation. Ces étapes sont réalisées majoritairement avec le package xcms (<http://bioconductor.org/packages/release/bioc/html/xcms.html>) (Smith *et al.* 2006) de R (<https://www.r-project.org/>) (R Core Team 2016). Les données finales se présentent sous la forme d'une matrice de données brutes comprenant 102 individus, des QC et un total de 3 378 ions.

2) *Contrôle de l'acquisition avec les QC*

Le bon déroulement de l'acquisition des données lors de l'analyse sur le spectromètre de masse est contrôlé via une analyse en composante principale (ACP) des données avec les QC. Deux traitements des données brutes sont nécessaires avant de réaliser cette analyse : la normalisation des données et le traitement des valeurs manquantes. Ici une normalisation par somme des intensités de tous les pics d'un individu est utilisée comme facteur de normalisation, chaque intensité est divisée par la somme totale des intensités d'un individu et multipliée par 100. Toutes les valeurs manquantes d'un ion sont remplacées par une petite valeur d'intensité qui est la plus petite valeur d'intensité de toute la matrice divisée par deux (Xia & Wishart 2011). Cette petite valeur correspond à une valeur du bruit de fond. D'après la Figure 45, les

QC obtenus dans le cadre de notre étude sont bien regroupés au centre des individus ce qui nous permet de valider la qualité de l'injection et nous autorise à poursuivre les analyses.

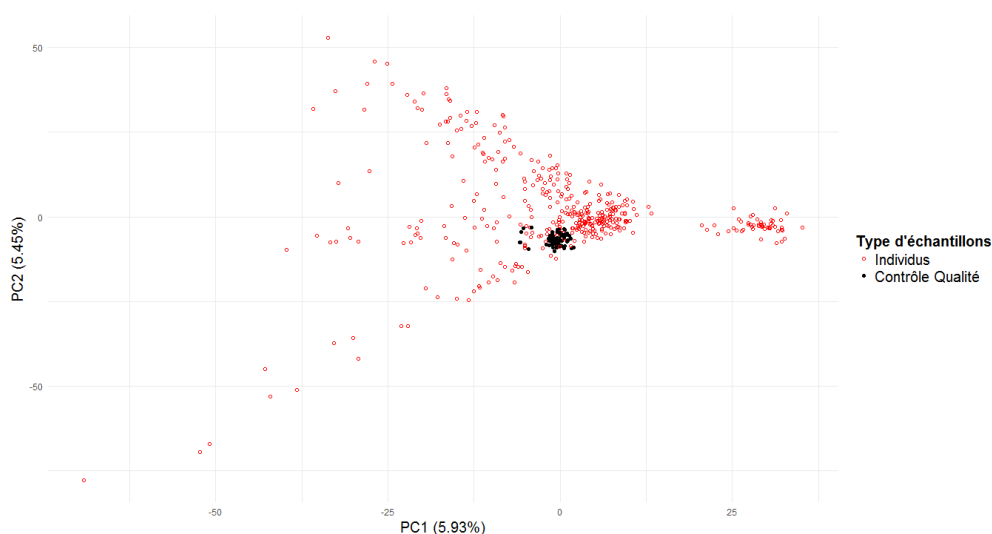


Figure 45 : Représentation graphique de l'Analyse en Composante principale (ACP) de toute la matrice de données de métabolomique.

3) Préparation de la matrice de données pour les analyses multivariées

Plusieurs étapes de traitement des données sont appliquées sur la matrice de données brutes :

- Les ions dont l'écart type relatif (RSD : *Relative Standard Deviation*, écart type divisé par la moyenne) parmi les QC est supérieure à 20% sont éliminés de toute la matrice de données, ce qui correspond à une perte de 68% des ions détectés. Cette étape permet d'éliminer les ions dont les valeurs d'intensité ne sont pas homogènes parmi les QC, alors qu'elles devraient l'être puisque le même QC est injecté tout au long de l'analyse. Ces ions ne sont pas fiables au sein des QC, ils sont donc enlevés de toute la matrice.
- L'histogramme des valeurs d'intensité de tous les QC (Figure 46) indique que les valeurs entre 0 et 1 000 sont les plus nombreuses et ainsi sont assimilées au bruit de fond. Le bruit de fond est la fluctuation du signal émis par l'instrument sans échantillon. Les ions dont l'intensité est inférieure à 10^3 sont considérés comme du bruit et sont désignés comme valeurs manquantes. Tous les ions ayant 100% de valeurs manquantes sont éliminés, ce qui correspond à une perte de 7% des ions après avoir enlevé les 68%.
- Les valeurs manquantes d'un ion sont ensuite remplacées par la moyenne des valeurs correspondantes de cet ion parmi les k (ici k=10) plus proches individus en terme de distance euclidienne des intensités à l'aide de l'algorithme K-nearest neighbour (Xia &

Wishart 2011; Hrydziusko & Viant 2012). Ainsi une valeur unique est utilisée pour remplacer toutes les valeurs manquantes d'un ion. La matrice de données ne contient ainsi plus aucune valeur manquante.

- Les données sont ensuite normalisées par PQN (*Probalistic Quotient Normalisation*) (Dieterle *et al.* 2006), cette méthode de normalisation permet de prendre en compte les changements de concentration des ions.
- Transformation Glog (Generalised logarithm) des données pour minimiser la variation (Parsons *et al.* 2007).
- Elimination des QC de la matrice

Ces prétraitements et transformations de la matrice brute de données initiales donne une matrice de données nettoyées comprenant 51 individus *Ca. femoratus*, 51 individus *Cr. levior* pour un total de 764 ions.

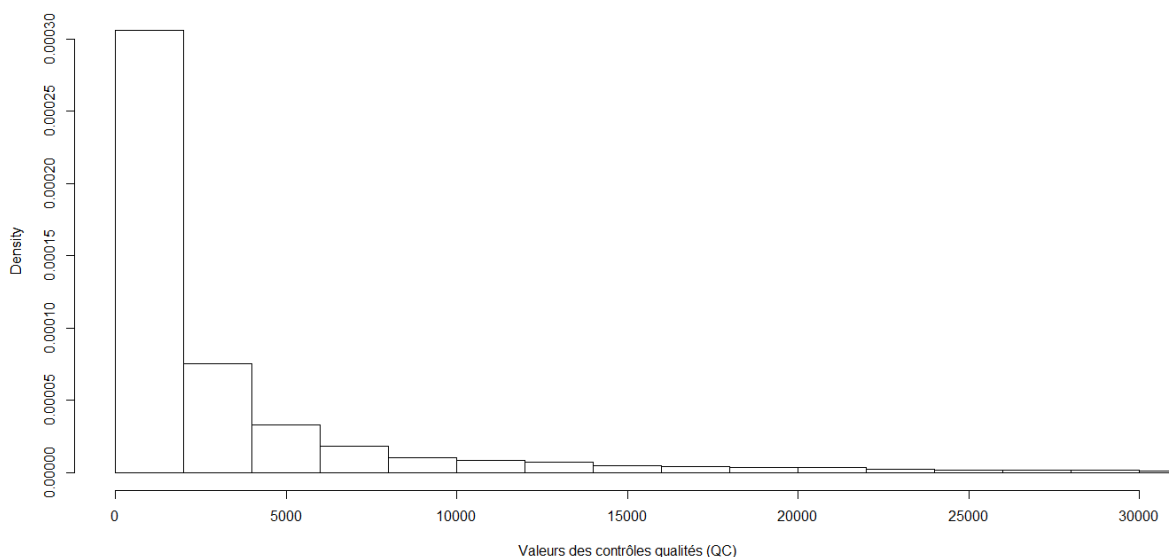


Figure 46 : Histogramme des valeurs d'intensité des ions de tous les QC de la matrice de données de métabolomique.

C. Séquençage haut débit du microbiote bactérien cuticulaire

1) *Extraction des ADN, amplification 16S et séquençage haut débit*

Les ADN des eaux de lavage sont extraits selon le protocole QIAamp que nous avons identifié comme étant le meilleur compromis dans l'obtention de données fiables sur les communautés bactériennes de la cuticule de fourmis (cf Chapitre 2). Brièvement, ce protocole sur colonne comprend deux étapes de lyse enzymatique successives (lysozyme et protéinase K), puis une capture des ADN sur une membrane de silice de porosité faible, suivi des étapes

de rinçage et d'élution de l'ADN. La région v5-v6 du gène ARNr 16S (295 bp) est utilisée comme marqueur de la diversité et de la composition des communautés bactériennes. Cette région est ensuite amplifiée par PCR, en duplicat. Les amorces en 5' et 3' pour l'amplification sont étiquetées avec des combinaisons de tags (fragments d'ADN de 8 nt de composition connue) permettant de réattribuer chaque séquence à son échantillon respectif après séquençage. Les amplicons tagués sont ensuite multiplexés et séquencés sur une plateforme de séquençage Illumina MiSeq avec la technologie paired-end 2x250 bp. Cette dernière étape a été réalisée en collaboration avec le laboratoire EDB et la plateforme Genotoul à Toulouse.

2) *Prétraitements bioinformatiques*

Ce travail a été réalisé en novembre 2016 par Lucie Zinger de l'Institut de Biologie de l'Ecole Normale Supérieure (IBENS) à Paris. Les étapes de traitement bio informatiques des données de séquençage sont similaires à celles détaillées dans la partie « *Sequence analysis* » du matériel et méthodes du chapitre 2. Ces étapes sont rappelées brièvement ci-dessous :

- L'assemblage des séquences en 5' et 3' pour former une séquence consensus.
- L'attribution des lectures aux échantillons grâce aux combinaisons de tags.
- La déréplication des lectures : les lectures ayant une séquence strictement identique sont regroupées. Seul la séquence et le nombre de lectures correspondant à chaque individu sont gardés. Ceci sert à réduire la taille des données.
- Les séquences présentant une similarité de séquence de 97% sont regroupées en OTUs. La séquence la plus abondante de l'OTU est retenue comme la séquence représentant l'OTU.
- Les OTUs sont assignés taxonomiquement en comparant la séquence représentative de l'OTU contre une banque de données de séquences spécifique aux gènes ARN ribosomal bactériens.
- Trois étapes de nettoyage se succèdent avec l'élimination des OTUs de mauvaise qualité (chimères ou séquences dégradées), des OTUs contaminant et les faux positifs résultant du phénomène de tag-switching.

A la fin de ce processus, les lectures des duplicatas de PCR sont sommées pour chaque individu car la variation entre les réplicats biologiques est largement plus importante et intéressante que celle entre réplicats de PCR. Les communautés bactériennes obtenues à partir des 60 individus de *Ca. femoratus*, des 60 individus de *Cr. levior* et des 24 échantillons de nids correspondent à 18 857 OTUs, correspondant à 907 508 lectures. Les taxons trop rares (i.e. <

20 reads sur l'ensemble du jeu de données) sont définis comme non fiables et sont exclus de l'analyse. Ceci diminue le nombre d'OTUs total de 18 857 à 2 453 et le nombre de lectures de 7,7% (907 508 reads à 838 248 reads). Les individus pour lesquels nous avons obtenu moins de 100 reads sont considérés comme des échantillons non fiables et sont exclus également (perte de 2 individus).

D. Analyses statistiques

1) *Richesse en OTUs et en ions*

Pour évaluer si les communautés bactériennes cuticulaires ont été échantillonnées de manière appropriée des courbes de raréfaction ont été tracées pour chaque échantillon. Si les communautés bactériennes n'ont pas été échantillonnées correctement, alors il est nécessaire de standardiser les données. Une méthode souvent employée pour standardiser la profondeur de séquençage entre les échantillons sur toute la table de comptage est de raréfier les données pour une taille d'échantillon choisie (ici 400 reads). La raréfaction s'effectue avec la fonction `rrarefy` du package R `vegan` (www.cran.r-project.org/web/packages/vegan/index.html). Elle consiste à réaliser un échantillonnage aléatoire sans remise de reads pour une taille d'échantillonnage donnée pour chaque élément de la table. La matrice "standardisée" sera utilisée pour la suite des analyses.

Les données de métabolomique ne sont pas des données de comptage, mais d'intensité. Afin d'évaluer et de comparer la richesse spécifique en ions des individus, les données sont transformées en présence/absence avec 0 et 1 avant l'étape de remplacement des valeurs manquantes selon l'algorithme KNN. Ainsi, toutes les valeurs manquantes (biologiquement non présente ou en dessous du seuil de détection de l'appareil) et toutes les valeurs assimilées comme du bruit (intensité inférieure à 10^3) sont remplacées par 0 et les valeurs d'intensité par 1 pour cette analyse de richesse.

La stratégie d'analyse statistique nécessaire à ces données de décompte d'OTUs et d'ions par individu est basée sur l'utilisation d'un modèle linéaire généralisé (*Generalised Linear Model*, GLM). Les données de comptages étant surdispersées, nous avons choisi de baser le modèle sur la loi binomiale négative. Ce test est réalisé avec l'aide de la fonction `glm.nb` du R package MASS (Ripley *et al.* 2015). L'effet des variables « sites de prélèvements », « espèces » ainsi que leur interaction ont été évalués avec une analyse de variance ANOVA de type II. Des comparaisons multiples sont réalisées avec le test post-hoc de Tukey, en corrigeant les p-values avec la méthode FDR (False Discovery Rate).

2) OTUs partagés entre les deux espèces

Afin de déterminer les taxons communs aux deux espèces de fourmis, seuls les OTUs détectés dans plus de 50% des répliques biologiques d'une espèce pour un site ont été considérés. Parmi ceux-ci, les OTUs partagés entre les deux espèces au sein d'un même site le sont lorsque l'OTU est détecté dans plus de 50% des répliques biologiques de la première espèce et dans plus de 50% des répliques biologiques de la seconde espèce du même site. Les OTUs partagés sont visualisés sur les diagrammes de Venn par les zones des cercles qui se chevauchent. Ces diagrammes de Venn ont été réalisés à l'aide du [R package VennDiagramm](#) (Hanbo Chen 2016).

3) Composition des communautés d'ions et d'OTUS

Les effets de site de prélèvement et de l'espèce sur la composition des OTUs et des ions cuticulaires sont étudiés premièrement par une analyse non supervisée, où l'on cherche à visualiser les dissimilarités entre individus. L'ordination se fait à partir d'une matrice de distance car à ce stade la question porte sur la dissimilarité globale entre les individus, plutôt que sur les variables. Les différences de composition des communautés sont calculées avec l'indice de Bray-Curtis sur des données préalablement transformées avec la transformation d'Hellinger pour les données microbiennes (Legendre & Legendre 2012) et sur des données préalablement normalisées et transformées en glog pour les données chimiques (Di Guida *et al.* 2016). La variance de la matrice de distance est visualisée avec une analyse en coordonnées principales (PCoA ; *Principal Coordinate Analysis*), qui s'apparente à l'ACP. Les effets du site de prélèvement, de l'espèce et de leur interaction sont testés avec une PERMANOVA (*Permutational Multivariate Analysis of Variance*). Ces analyses sont conduites avec le package R *vegan* (Anderson 2001).

Dans un deuxième temps, nous utilisons une analyse de type supervisée. L'analyse discriminante par la méthode des moindres carrés partiels (*Partial Least Squares/Projection to Latent Structures-Discriminant Analysis*, PLS-DA) est une régression classique PLS où la variable de réponse est catégorielle, ici ce sont les classes des individus comme l'espèce ou le site. PLS-DA est une méthode de classification dont le principe est de maximiser la covariance entre la matrice des variables et la matrice des classes. L'analyse supervisée de type sPLS-DA (*sparse PLS DA*) permet de sélectionner les variables expliquant les classes dans l'analyse de la composition des communautés par rapport à ces différentes classes connues. La sPLS-DA utilise la pénalisation LASSO pour sélectionner les variables (OTUs ou ions) les plus discriminatives du modèle PLS-DA (Lê Cao *et al.* 2011). La sPLS-DA est un modèle prédictif qui sélectionne les OTUs dans chaque composant en construisant une règle de classification.

Une étape d'ajustement ou de « *tuning* » est effectuée pour sélectionner le plus petit nombre de variables nécessaires à chaque composant du modèle sPLS-DA tout en conservant un taux d'erreur de classification le plus bas possible.

Ici, les différences de composition des communautés sont déterminées à partir des données préalablement normalisées CSS (*Cumulative Sum Scaling normalisation*) (Paulson *et al.* 2013) puis transformées log pour les OTUs et sur des données préalablement normalisées et transformées (glog) pour les ions (Di Guida *et al.* 2016). L'interprétation des résultats d'une analyse sPLS-DA repose en premier lieu sur la validation du modèle. En effet, les représentations graphiques et les séparations que l'on peut y observer ne reflètent pas obligatoirement la différence qui existe réellement entre les classes. En revanche, si le modèle est correctement validé, la séparation observée est significative et représente une réelle différence dans les données. La performance de classification du modèle sPLS-DA final peut être évaluée à l'aide de la fonction `perf` du package `mixOmics`. Les taux d'erreur de classification moyens par rapport au nombre de variables sélectionnées sont obtenus par validation croisée (50 répétitions).

Pour préciser la fonction des métabolites détectés, nous avons fait une analyse du défaut de masse de Kendrick. Cette analyse permet d'observer parmi une liste d'ions si certains de ces ions se regroupent en famille chimique. Le défaut de masse est défini comme la différence entre la masse exacte d'un composé et sa masse nominale. Les ions correspondant aux molécules d'une même famille, par exemple les ions d'une série d'alcane ont le défaut de masse de Kendrick identiques mais diffèrent de part leur masse nominale. Ils se retrouvent ainsi alignés dans la représentation graphique du défaut de masse de Kendrick (masse nominale en fonction de défaut de masse) (Kendrick 1963) (Hughey *et al.* 2001).

III. Résultats et discussions

A. Richesse en OTUs et en ions

Les courbes de raréfaction pour chaque échantillon permettent de visualiser si la diversité bactérienne est suffisamment couverte pour tous les échantillons, indépendamment du facteur espèce/nids et du facteur site. Les courbes n'atteignent pas un plateau pour les deux espèces de fourmis, ce qui signifie que les communautés ne sont pas suffisamment échantillonnées. Les courbes de raréfaction des nids sont beaucoup plus courtes et loin d'être saturées par rapport à celles des deux espèces de fourmis (Figure 47). Ceci est confirmé avec une analyse de variance faite sur les individus, où le facteur espèce/nid a un effet significatif sur le nombre de lectures (ANOVA ; $X^2=14.34$, $df=2$; $p<0,001$). Les échantillons de sols des Nouragues et certains de Petit Saut ont significativement moins de reads que les individus fourmis ou les échantillons des sols de Kaw d'après les comparaisons multiples.

En utilisant des données raréfiées pour corriger ces différences d'échantillonnage, ni l'espèce et ni le site de prélèvement ont un effet significatif sur le nombre d'OTUs (ANOVA : espèces : $X^2=0,054$; $df=2$; $p=0,973$ et site : $X^2=0,34$; $df=2$; $p=0,844$) (Figure 47).

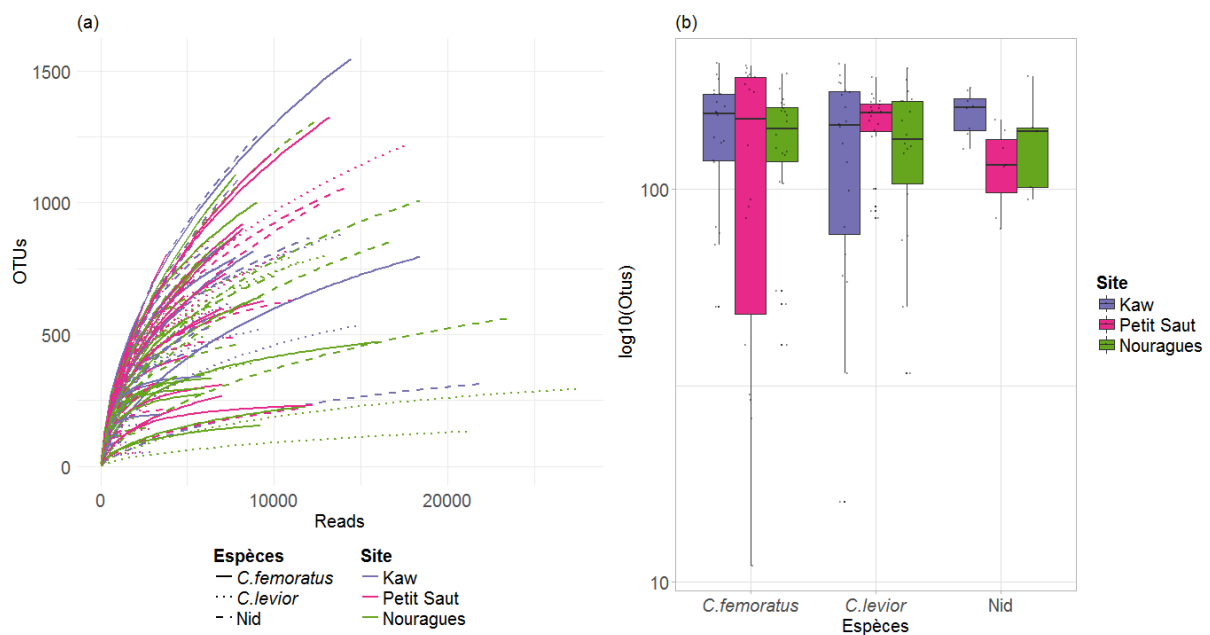


Figure 47 : Courbe de raréfaction (a) et richesse (nombre d'OTUs) à partir de données raréfiées (b) pour *Camponotus femoratus*, *Crematogaster levior* et les nids prélevés sur trois sites de prélèvement : à la montagne de Kaw (violet), à Petit Saut (rose) et à la station scientifique des Nouragues (vert).

Il n'y a pas de différence significative du nombre d'OTUs entre les cuticules des deux espèces de fourmis et les échantillons de nids. La nature particulière de la cuticule composée principalement de chitine et la présence de glandes excrétrices (notamment glande métapleurale) aurait pu laisser suggérer que l'on détecte moins d'OTUs que dans les échantillons de nids. Contrairement à ce que nous supposions, la cuticule n'apparaît pas comme un filtre de sélection au niveau de la richesse en OTUs. Néanmoins, il est possible que les communautés des nids restent sous échantillonnées malgré la raréfaction. Ces échantillons sont riches en inhibiteurs de PCR (matière organique), la raréfaction portant sur les séquences, elle ne permet pas de corriger un sous-échantillonnage ayant eu lieu à l'étape de la PCR.

En comparant le nombre d'ions entre les individus, le facteur espèce a un effet significatif et moins important (ANOVA : $X^2=4,3$, $df=1$; $p=0,04$) que le site de prélèvement (ANOVA : $X^2=19,47$; $df=2$; $p<0,001$) (Figure 48). Les individus *Cr. levior* des Nouragues ont significativement plus d'ions que les *Cr. levior* de Petit-Saut et de Kaw d'après les comparaisons multiples (toutes $p<<0,001$). En revanche, il n'y a pas de différence en termes de richesse d'ions entre les deux espèces du jardin de fourmis.

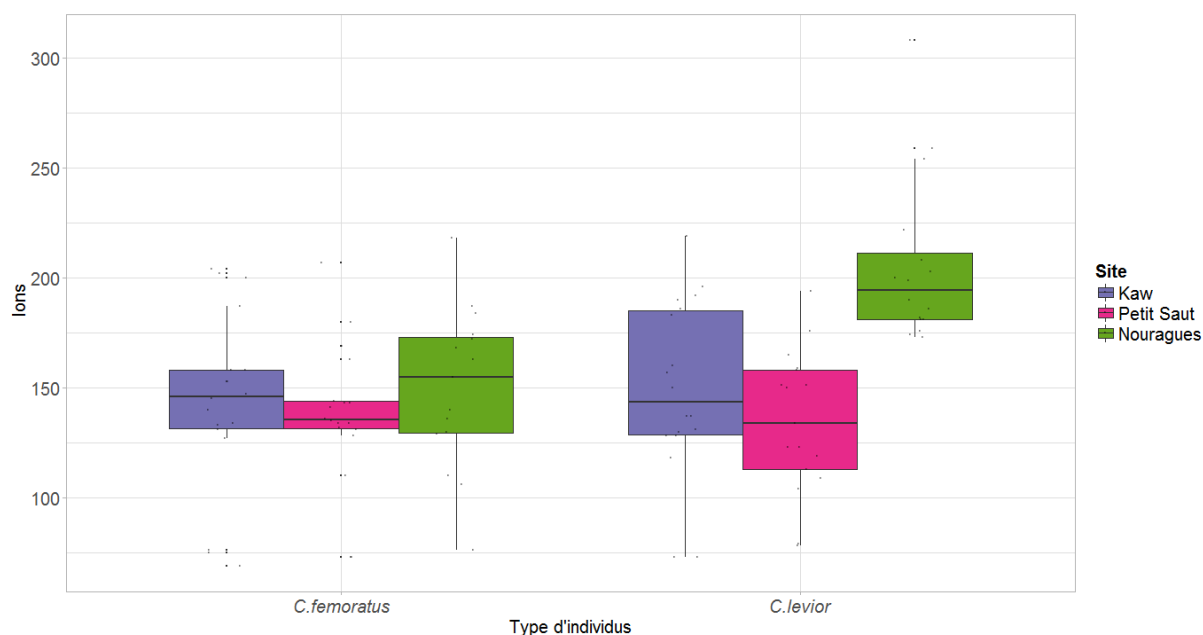


Figure 48 : Observation des données brutes de métabolomique représenté par le nombre d'ions pour *Camponotus femoratus* et *Crematogaster levior* sur trois sites de prélèvement : à la montagne de Kaw (violet), à Petit Saut (rose) et à la station scientifique des Nouragues (vert).

B. OTUs partagés par les deux espèces

1) Résultats

L'étude de la parabiose entre *Ca. femoratus* et *Cr. levior* permet d'accéder aux OTUs partagés par les deux espèces au sein du jardin de fourmi. Le nombre d'OTUs partagé entre les deux espèces de fourmis, incluant des OTUs partagés avec le nid pour certains, est de 51 à Kaw, 16 à Petit Saut et 27 aux Nouragues (Figure 49). Parmi ces OTUs, seul 8 d'entre eux sont partagés par les deux espèces sur les trois sites (cf Annexe 7). Ces 8 OTUs partagés par les deux espèces sur les 3 sites sont une Acidobacteria (*Acidobacteria* sp.), trois Alphaproteobacteria (*Bradyrhizobium* sp., *Variibacter* sp. et *Burkholderiales* sp.) et quatre Actinobacteria (*Microbacterium* sp., *Nakamurella* sp., *Mycobacterium* sp. et *Jatrophihabitans* sp.). Notons qu'aucun de ces 8 OTUs n'est totalement absent des échantillons de nid.

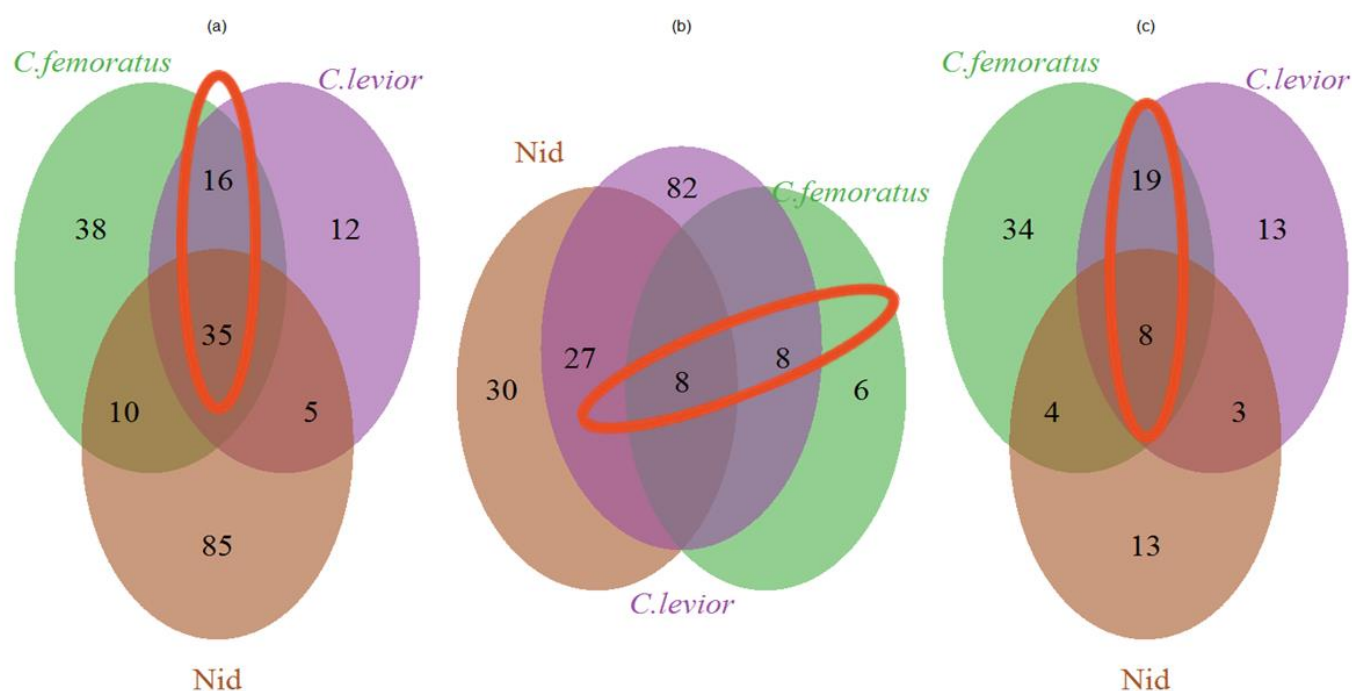


Figure 49 : Diagramme de Venn des OTUs pour les deux espèces de fourmis *Camponotus femoratus* et *Crematogaster levior* ainsi que leur nid à Kaw (a), à Petit Saut (b) et aux Nouragues (c).

1) Discussions

Les deux espèces du jardin de fourmis ont 8 OTUs en commun, et ces 8 OTUs sont retrouvés sur les fourmis des jardins des trois sites de prélèvement. Il est peu probable que ces OTUs soient des contaminants puisque nous avons consciencieusement enlevé les contaminations potentielles du jeu de données et que ces OTUs sont affiliés à des clades qui ne sont pas des contaminants notoires. Ce résultat ne nous permet pas de savoir si ces OTUs sont constitutifs de la parabiose, i.e. partagés par les deux espèces via une transmission horizontale liée à la parabiose. Plus exactement, s'ils sont transmis verticalement pour les deux espèces, ou

pour une espèce et acquis horizontalement par contact avec l'autre, ou s'ils ont été acquis dans l'environnement par chacune des deux espèces. Si ces OTUs sont impliqués dans la parabiose, il est également possible que les hôtes contaminent leur nid avec ces symbiontes. Une étude plus poussée des abondances de ces OTUs dans les différents échantillons pourrait permettre de tirer des conclusions plus robustes.

Si l'on s'intéresse à l'identité et au rôle potentiel de ces OTUs partagés par les deux espèces, l'hypothèse de bactéries impliquées dans la parabiose est envisageable. En effet, la présence d'actinobactéries connues pour leur production en métabolites secondaires bioactifs pourrait permettre aux hôtes de se défendre contre des infections entomopathogènes au sein du jardin. Les actinobactéries des genres *Microbacterium*, *Nakamurella*, et *Jatrophihabitans* sont des genres connus pour leur production en métabolites bioactifs (Barka *et al.* 2016). De plus, ces OTUs pourraient être partagés dans un but de communication : les raisons de la tolérance des colonies de ces deux espèces de fourmis l'une envers l'autre dans un même habitat ne sont pas connues. Bien que peu étudié, il a été démontré que les métabolites bactériens volatiles de la cuticule des fourmis peuvent avoir une influence sur la reconnaissance entre les individus (Davis *et al.* 2013). Une équipe a testé expérimentalement si les micro-organismes contribuent effectivement à la reconnaissance entre les individus d'un même nid chez les fourmis *Pogonomyrmex barbatus*. Au cours de cette expérience, le microbiote externe des fourmis a été modifié par l'application topique d'antibiotiques ou de cultures microbiennes. Les résultats montrent que les fourmis dont le microbiote externe a été augmenté avec des cultures microbiennes ont été rejetées (Dosmann *et al.* 2016). Dans le cas des jardins de fourmis, nous pourrions donc tester l'hypothèse de l'implication du microbiote dans la parabiose en traitant la cuticule d'individus des deux espèces par des antibiotiques et d'observer si les individus traités sont toujours acceptés par les colonies des deux espèces.

C. Composition des communautés d'OTUs

Comparer l'importance relative des facteurs « espèce » et « site de prélèvement » des individus sur la composition des communautés bactériennes peut apporter des éléments de réponse quant à l'acquisition de ces communautés bactériennes. Dans un premier temps, et par rapport à nos hypothèses, il est nécessaire de confirmer que la composition des communautés bactériennes est différente entre les trois sites de prélèvements. Deuxièmement, nous pourrions tester si le facteur site est plus important que le facteur espèce sur la composition du microbiote cuticulaire.

1) Analyses non supervisées

(a) Communautés bactériennes des sites

La PCoA comprenant les échantillons du nid présente une forte disparité en trois groupes correspondant aux trois sites (Figure 50). Ces observations sont confirmées statistiquement par l'analyse de PERMANOVA (site de prélèvement : $F_{2,20}=5,19$; $p < 0,001$; $R^2=0,36$). Le site de prélèvement explique 36% de la variance observée entre les différents échantillons de nids. Les communautés bactériennes sont différentes selon les trois sites de prélèvement.

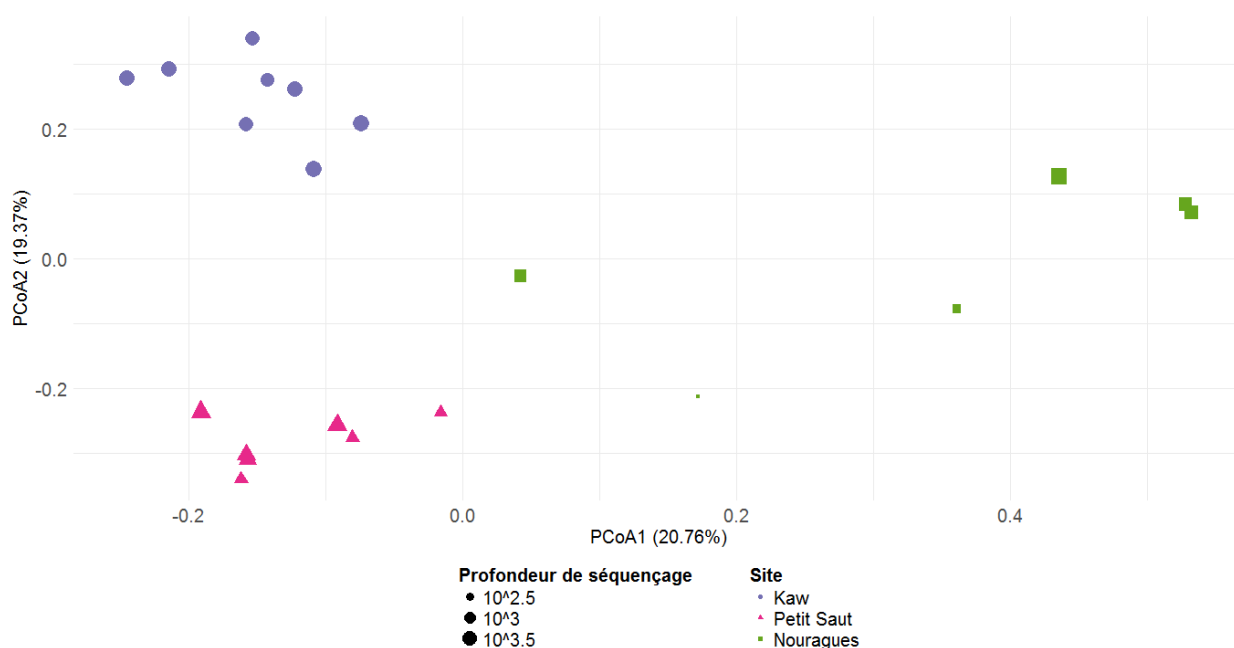


Figure 50 : Similarités de composition des communautés d'OTUs pour tous les échantillons de nids. La taille des points représente la profondeur de séquençage (nombre de reads) pour chaque échantillon.

Les échantillons des nids sont un ensemble représentatif de l'environnement direct de la fourmi dans le jardin, c'est-à-dire des racines et de la terre. Les fourmis apportent la terre au jardin à partir de la terre du sol environnant. Les communautés bactériennes du sol sont fortement influencées par le pH, l'eau et les teneurs en carbone ou en azote (Hanson *et al.* 2012) qui diffèrent probablement entre les différents sites d'étude. De plus, la composition taxonomique des plantes constituant le jardin pourrait aussi avoir un rôle sur la composition des communautés bactériennes de la rhizosphère (i.e. interface sol/racine) via leurs diversités et la

composition des exsudats racinaires (Nannipieri *et al.* 2008). Nos sites d'échantillonnages présentent donc bien des pools d'espèces microbiennes différentes.

(b) Microbiote bactérien des cuticules

La PCoA comprenant les deux espèces de fourmis présente des individus dispersés en regard du facteur espèce, mais permet de mettre en évidence un groupe dont tous les individus sont issus du site de prélèvement des Nouragues (Figure 51). Ces observations sont confirmées statistiquement par l'analyse de PERMANOVA (espèces : $F_{1,117}=4,92$; $p<<0,001$; $R^2=0,036$ site de prélèvement : $F_{2,117}=6,96$; $p<<0,001$; $R^2=0,10$, interaction : $F_{2,117}=2,51$; $p<<0,001$; $R^2=0,037$). Le facteur espèce explique seulement 3,6% de la variance observée, ce qui est inférieur au facteur site de prélèvement de 10%. Ces éléments suggèrent donc que la plupart des OTUs et/ou les OTUs les plus abondants des microbiotes cuticulaires ne sont pas spécifiques de l'espèce et qu'ils sont acquis dans l'environnement.

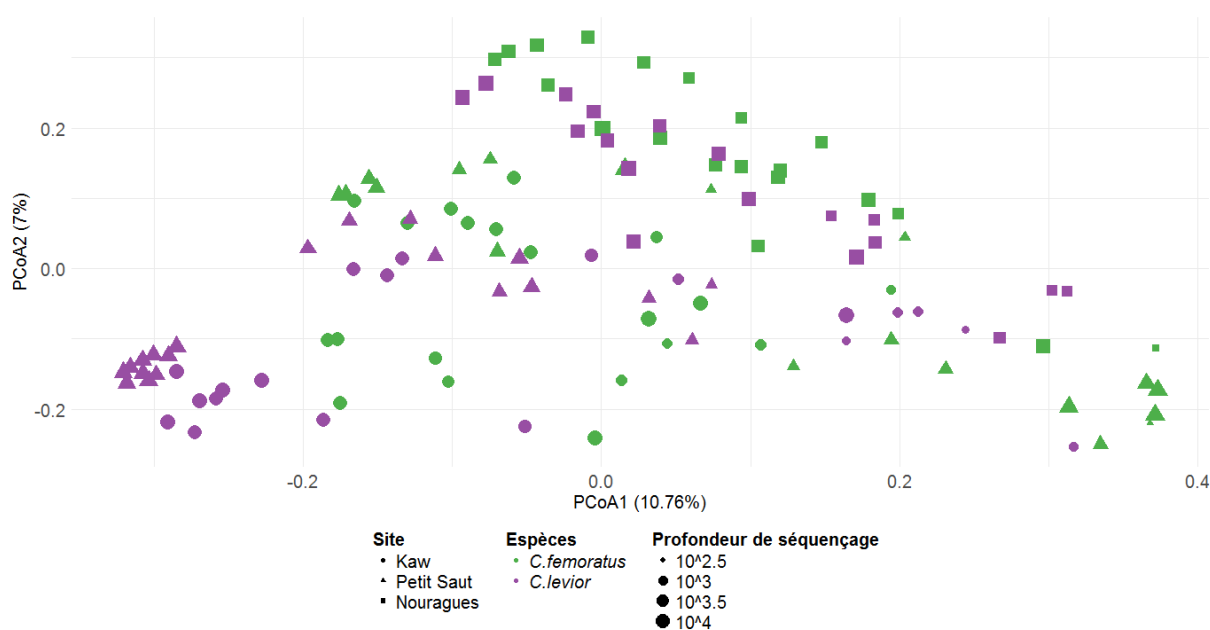


Figure 51 : Similarités de composition des communautés d'OTUs pour tous les individus. La taille des points représente la profondeur de séquençage (nombre de reads) pour chaque individu.

2) Analyses supervisées

D'après les analyses non supervisées sur la composition des communautés d'OTUs, l'information de l'espèce n'est pas le facteur expliquant le plus les dissimilarités observées entre les communautés bactériennes de la cuticule des deux espèces de fourmis habitant le jardin de

fourmis. Néanmoins, il explique une faible part de la variance observée (3,6%) : quels sont les OTUs différents entre les deux espèces ? Appartiennent-ils aux *Ca. femoratus* ou aux *Cr. levior* ?

(a) sPLS-DA et espèces

Le modèle de sPLS-DA avec la variable catégorielle espèce est validé, le taux d'erreur de classification est déjà faible sur le premier composant (16%), et il continue de diminuer avec l'ajout du deuxième composant (13%). Le taux d'erreur diminue avec l'ajout des composants, lorsqu'il ne diminue plus c'est qu'il n'est plus nécessaire d'ajouter des composants. L'étape de « tuning » sur ce modèle sPLS-DA propose de conserver 10 OTUs seulement pour le premier axe et 100 OTUs pour le deuxième axe (Figure 52).

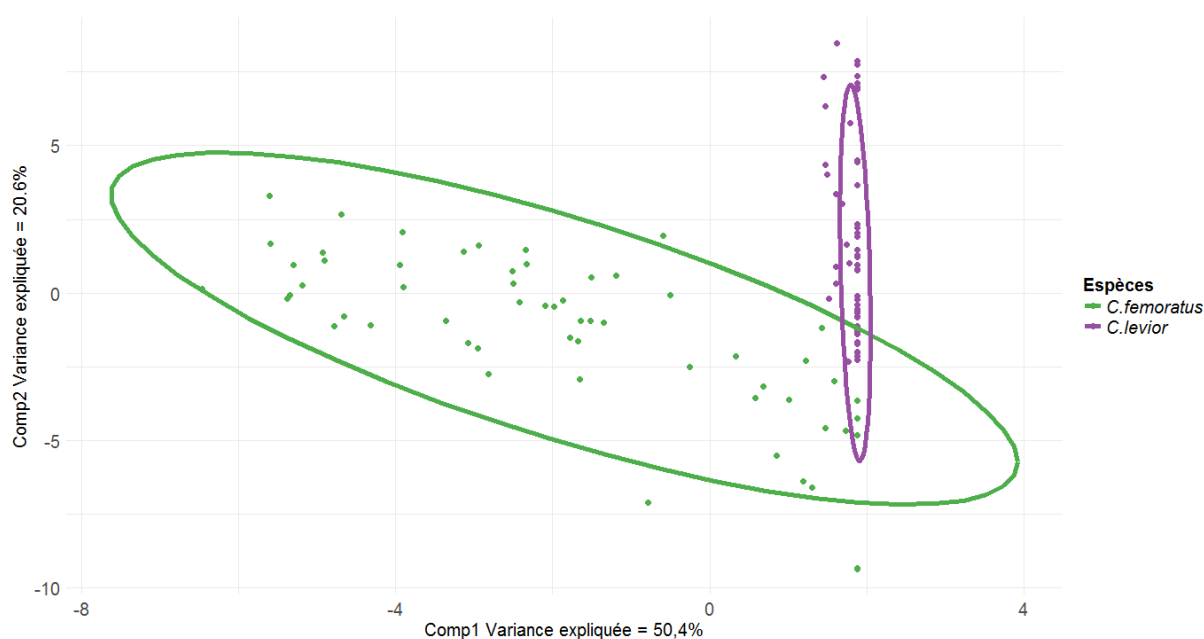


Figure 52 : sPLS-DA avec la variable catégorielle espèce.

Ici, la sPLS-DA sélectionne les OTUs les plus discriminants qui caractérisent le mieux chaque espèce. La Figure 53 est une représentation graphique de leur pouvoir discriminant, du poids de chaque variable dans l'analyse. Seulement 10 OTUs sont sélectionnés statistiquement par l'analyse de sPLS-DA pour différencier les *Ca. femoratus* des *Cr. levior*. Le petit nombre d'OTUs sélectionnés par cette analyse de sPLS-DA montre que l'espèce n'est pas le facteur discriminant à l'échelle de la communauté bactérienne. Néanmoins, ces OTUs pourraient être ceux qui sont transmis verticalement. Ceci confirme donc nos hypothèses de départ. L'ensemble de ces 10 OTUs sont des OTUs dont l'abondance est maximale chez les *Ca. femoratus* (Figure 53), suggérant donc que seule cette espèce comporte dans son microbiote des

OTUs transmis verticalement. Malheureusement, l'identification de ces OTUs est d'une faible résolution taxonomique, probablement parce que ces OTUs n'ont jamais été référencés auparavant, et il est donc difficile de savoir quelle est la fonction de ces bactéries pour leur hôte (Table 10).

Table 10 : Liste des 10 OTUs sélectionnés par le composant 1 du modèle sPLS-DA selon la variable catégorielle espèce classés par ordre d'importance décroissant.

Code	Domaine	Phylum	Classe	Ordre
seq_60406	Bacteria	-	-	-
seq_58250	Bacteria	Proteobacteria	-	-
seq_31107	Bacteria	Proteobacteria	-	-
seq_377057	Bacteria	-	-	-
seq_63142	Bacteria	-	-	-
seq_3564	Bacteria	Proteobacteria	Gammaproteobacteria	-
seq_27912	Bacteria	Actinobacteria	Actinobacteria	Actinomycetales
seq_22552	Bacteria	-	-	-
seq_79823	Bacteria	Proteobacteria	-	-
seq_18221	Bacteria	Verrucomicrobia	-	-

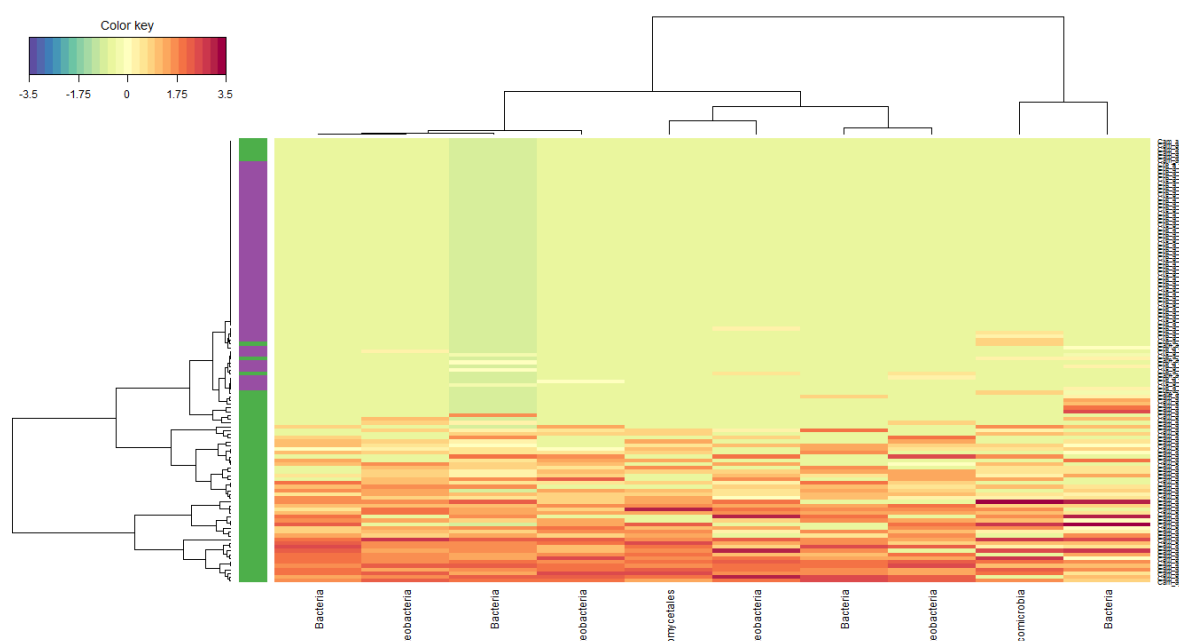


Figure 53 : Clustered image map, représentation graphique des 10 OTUs sélectionnés par le composant 1 du modèle sPLS-DA selon la variable catégorielle espèce (*Camponotus femoratus* en vert et *Crematogaster levior* en violet).

(b) sPLS-DA et site de prélèvement

L'information du site de prélèvement (Kaw, Petit Saut et les Nouragues) est le facteur expliquant le plus les dissimilarités observées entre les communautés bactériennes de la cuticule des deux espèces de fourmis habitant le jardin de fourmis. Quel est le site de prélèvement contribuant le plus à ces dissimilarités observées ? Quels sont les grands phylums d'OTUs les plus représentés parmi ces OTUs spécifiques au site ?

Le modèle de sPLS-DA avec la variable catégorielle site de prélèvement est validé, le taux d'erreur de classification est important sur le premier composant (36%), et il diminue

fortement avec l'ajout du deuxième composant (13%). L'étape de « *tuning* » sur ce modèle sPLS-DA propose de conserver 100 OTUs pour le premier axe et 90 OTUs pour le deuxième axe (Figure 54). Ici, la sPLS-DA sélectionne les OTUs les plus discriminants qui caractérisent le mieux chaque site de prélèvement.

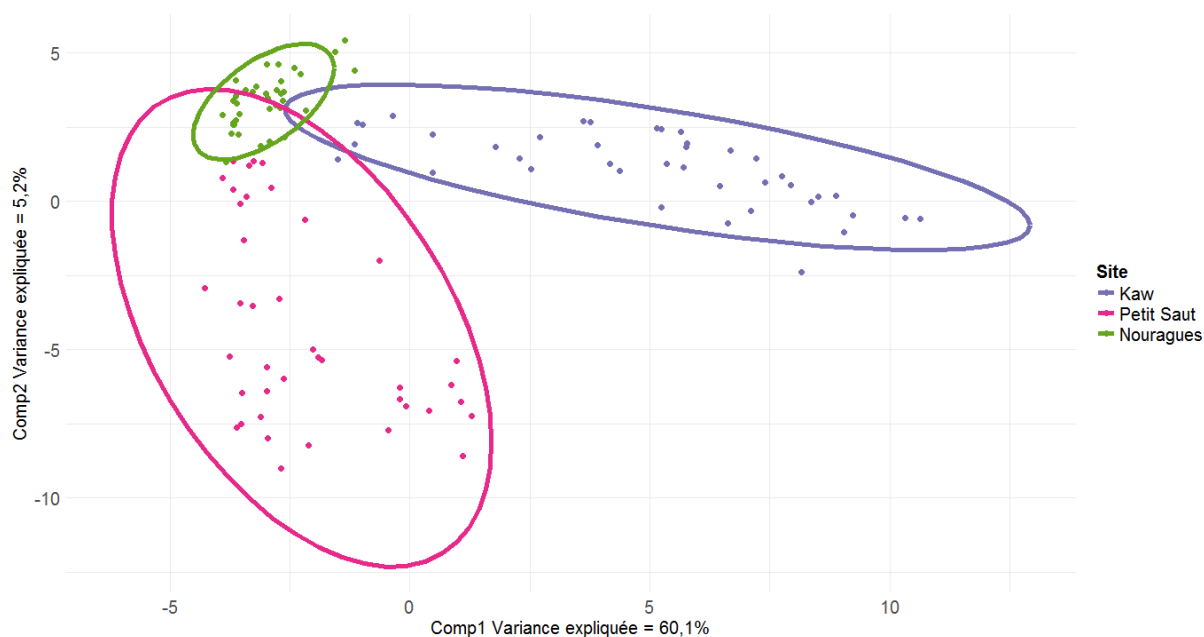


Figure 54 : sPLS-DA avec la variable catégorielle site de prélèvement.

La Figure 55 est une représentation graphique de leur pouvoir discriminant, la couleur des plots de contribution indique le site de prélèvement où il y a la moyenne la plus importante pour chaque OTU sélectionné. Plus particulièrement, les graphiques montrent l'abondance de chaque OTUs (grande abondance = grande valeur absolue) et dans quel site de prélèvement ils sont les plus abondants pour chaque composante sPLS-DA. Ce sont les 100 premiers OTUs qui séparent le mieux les 3 sites de prélèvement, peu importe l'espèce, et ce sont des OTUs majoritairement présents à Kaw. Puis le deuxième composant lui discrimine les individus selon le site de prélèvement surtout grâce à des OTUs spécifiques au site de Petit Saut. Les phylums majoritaires représentés parmi les OTUs sélectionnés par le premier composant ou le deuxième composant sont les Acidobacteria, Actinobacteria, Proteobacteria, Bacteroidetes et les Planctomycetes (résultats non présentés).

Bien qu'identifier le site de prélèvement contribuant le plus aux dissimilarités observées n'apporte pas d'éléments essentiels à nos interrogations principales sur l'acquisition du microbiote, nous pouvons tout de même conclure que le site de Kaw apparaît comme un milieu plus diversifié que le site de Petit Saut puis celui des Nouragues. Une hypothèse permettant

d'expliquer ces résultats, serait que les sites de la montagne de Kaw et de Petit Saut sont des sites plus proches de milieux anthropisés et seraient peut-être plus perturbés que le site des Nouragues. Ces perturbations pourraient être à l'origine des différences de composition des communautés observées entre les sites. Néanmoins, cette hypothèse reste à vérifier.

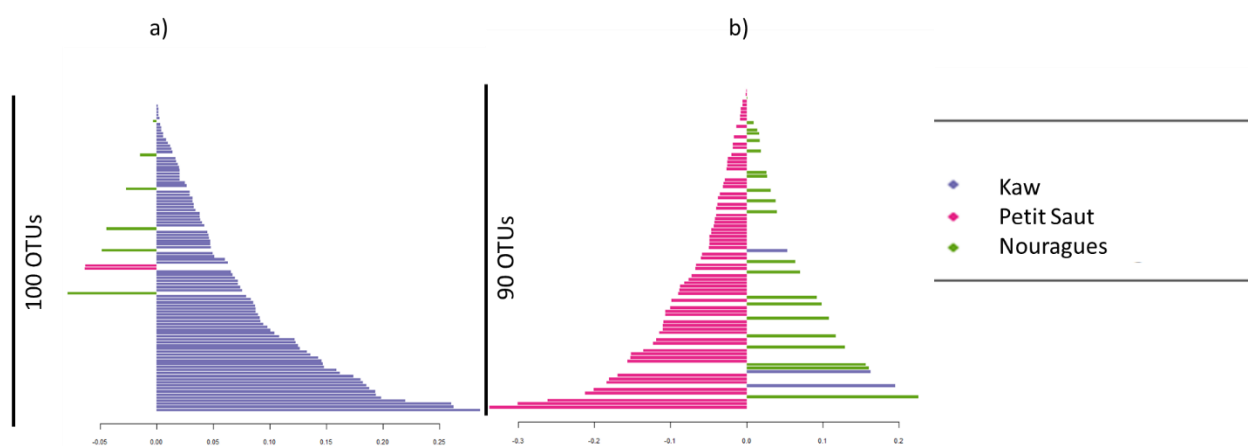


Figure 55 : Contribution plot des OTUs sélectionnés par le composant 1 (a)) et par le composant 2 (b)) de la sPLS-DA selon la variable catégorielle du site de prélèvement (Kaw en bleu, Petit Saut en rose et les Nouragues en vert).

D. Composition des communautés d'ions

L'analyse des différences des compositions d'ions entre ces deux espèces de fourmis et l'implication des facteurs comme l'espèce et le site de prélèvement des individus sur la composition d'ion peuvent apporter des éléments de réponse quant à l'origine de ces ions. Sont-ils issus de la fourmi ou de son microbiote cuticulaire ?

1) Analyses non supervisées

La PCoA présente deux groupes d'individus qui sont séparés selon l'axe 1. Ces deux groupes correspondent chacun aux deux espèces. L'axe 2 permet de discriminer un groupe dont tous les individus sont issus du site de prélèvement des Nouragues. Ces observations en deux dimensions sont confirmées statistiquement par l'analyse de PERMANOVA (effet de l'espèce : $F_{1,101}=23,174$; $p<<0,001$; $R^2=0,17$; effet du site de prélèvement : $F_{2,101}=5,467$; $p<<0,001$; $R^2=0,08$; interaction : $F_{2,101}=4,96$; $p<<0,001$; $R^2=0,07$).



Figure 56 : Similarités de composition des communautés d'ions pour les individus des deux espèces *Camponotus femoratus* et *Crematogaster levior*.

Ces résultats montrent que le facteur espèce est celui qui explique la plus grande part de la variabilité (17%) des communautés d'ions retrouvées sur les individus. Ces métabolites spécifiques à l'espèce de fourmi sont produits soit par la fourmi soit par le microbiote cuticulaire. D'autre part, 8% de la variabilité des communautés d'ions entre les individus s'explique par le site de prélèvement des individus. Etant donné l'importance du facteur site dans la variabilité des communautés bactériennes (Figure 51), l'hypothèse que les ions varient entre sites soient issus des bactéries spécifiques au site de prélèvement n'est donc pas rejetée. Toutefois, il est possible que le métabolome produit par une espèce de fourmi puisse varier selon différentes zones géographiques et ainsi constituer le chémotype d'une colonie, comme c'est le cas pour les CHCs (Sturgis, S.J. & Gordon, D.M. 2012). A ce stade de l'analyse, nos résultats ne permettent donc pas d'identifier si cette part de variabilité des ions liée au site de prélèvement est due à la fourmi elle-même ou à ses bactéries cuticulaires.

2) Analyses supervisées :

(a) sPLS-DA et espèces

Le modèle de sPLS-DA avec la variable catégorielle espèce est validé, le taux d'erreur de classification est faible sur le premier composant (13%) et reste stable sur le deuxième composant (13%). L'étape de « tuning » sur ce modèle sPLS-DA propose de conserver 100 ions pour le premier axe et 80 ions pour le deuxième axe (Figure 57). Ici, la sPLS-DA

sélectionne les ions qui caractérisent le mieux chaque espèce indépendamment du site de prélèvement.

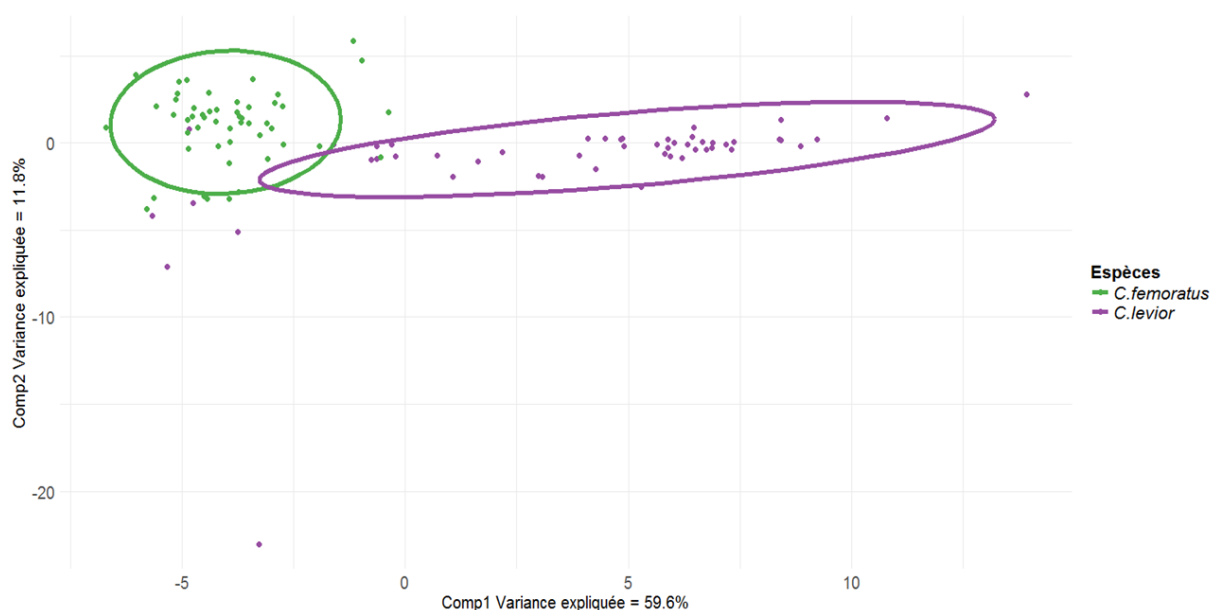


Figure 57 : sPLS-DA avec la variable catégorielle espèce.

La Figure 58 est une représentation graphique du pouvoir discriminant des ions pour séparer les deux espèces. La taille des plots indique l'importance de la contribution de l'ion pour l'espèce à laquelle il appartient. D'après cette figure, nous constatons que l'espèce *Ca. femoratus* a sur sa cuticule les métabolites les plus discriminants pour expliquer les différences principales entre les deux métabolomes. *Camponotus femoratus* est également l'espèce pour laquelle certains OTUs sont supposés être transmis verticalement (Figure 53). Ainsi, une hypothèse serait que ces ions sont corrélés aux OTUs potentiellement transmis verticalement chez l'espèce *Ca. femoratus*. Mais cette hypothèse reste à vérifier.

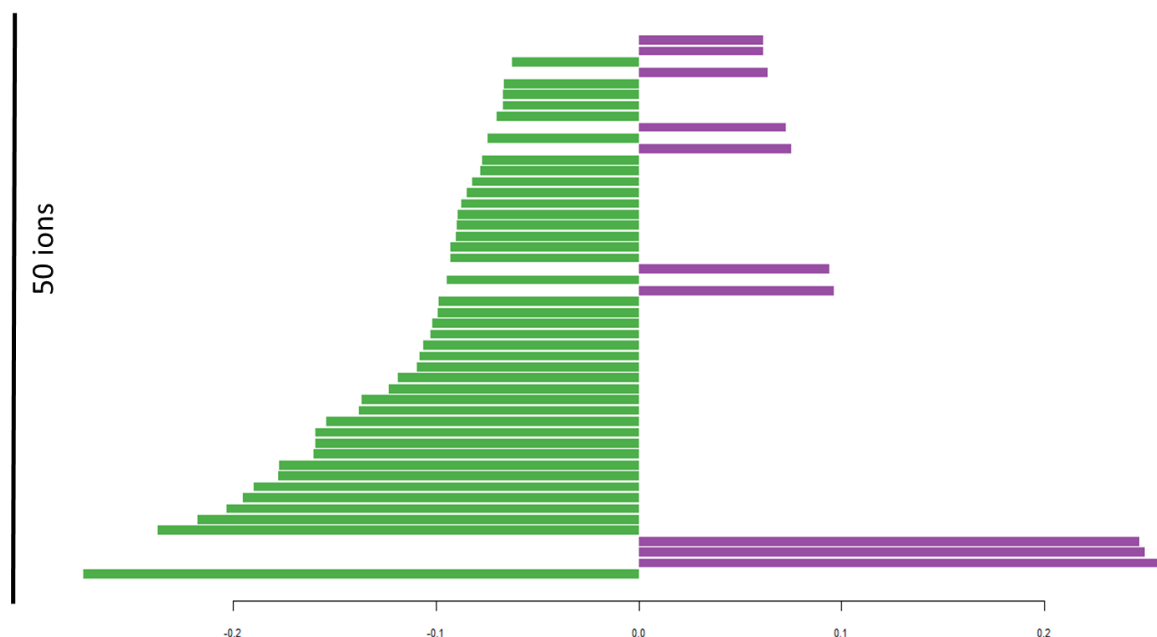


Figure 58 : Contribution plot des 50 premiers ions sélectionnés par le composant 1 de la sPLS-DA selon la variable catégorielle espèce (*Camponotus femoratus* en vert et *Crematogaster levior* en violet).

L'analyse de défaut de masse de Kendrick des 50 premiers ions sélectionnés par la composante 1 de cette analyse sPLS-DA ne montre pas de net alignement d'ions correspondant à une famille chimique particulière (Figure 59). D'une manière générale, on constate que ces ions possèdent une large gamme de poids moléculaire variant de 100 à 1100 m/z. La liste des 50 premiers ions a été comparée à la base de données Metlin (<https://metlin.scripps.edu>) en mode positif en autorisant les identifications possibles avec un delta ppm (taux d'erreur) inférieur à 10. Peu d'ions ont été identifiés par cette base de données. Ceci n'est pas surprenant car de nombreux métabolites bactériens et les métabolites de surface de fourmis tropicales n'ont pas fait l'objet d'étude en HRMS jusqu'à présent.

Les identifications les plus probables correspondent à des ions de la famille des phospholipides, des triglycérides (Table 11). D'après l'analyse de défaut de masse de Kendrick, d'autres ions peuvent s'aligner avec les ions identifiés suggérant qu'ils seraient des molécules de la même famille chimique (Figure 59). Ces ions sont vraisemblablement produits par les fourmis. En effet rappelons que chez les insectes les CHCs sont biosynthétisés dans les cellules spécialisées appelées oenocytes localisées sous la cuticule dans lesquelles la formation des chaînes hydrocarbonées est assurée par plusieurs enzymes telles que l'acétyl-CoA, les

déshydrogénases et les décarboxylases (Qiu *et al.* 2012; R & Sj 2015). Une fois produit, ces CHCs sont pris en charge par des transporteurs connus chez les insectes, les lipoprotéines. Ces dernières, associées à des phospholipides et à des triglycérides, s'assemblent sous forme de micelles contenant les CHCs et assure leur transport à la surface de la cuticule (Lockey 1988). Bien que la plupart des CHCs aient été étudiés, les transporteurs associés sont moins connus. Nos résultats suggèrent qu'il existe au même titre que les CHCs des différences de composition des molécules de transport chez les fourmis, mais nos résultats ne permettent pas, à partir de deux espèces seulement, de conclure sur l'utilisation de ces métabolites pour discriminer les espèces dans des études chimiotaxonomiques comme c'est le cas avec les CHCs.

Il est difficile d'identifier d'autres composés en l'absence de références dans les bases de données. L'une des solutions à l'avenir pour améliorer l'identification de ces ions est d'effectuer une analyse MS/MS afin de comparer non pas seulement le rapport m/z de l'ion mais également son profil de fragmentation utile pour réaliser des réseaux moléculaires (Wang *et al.* 2016).

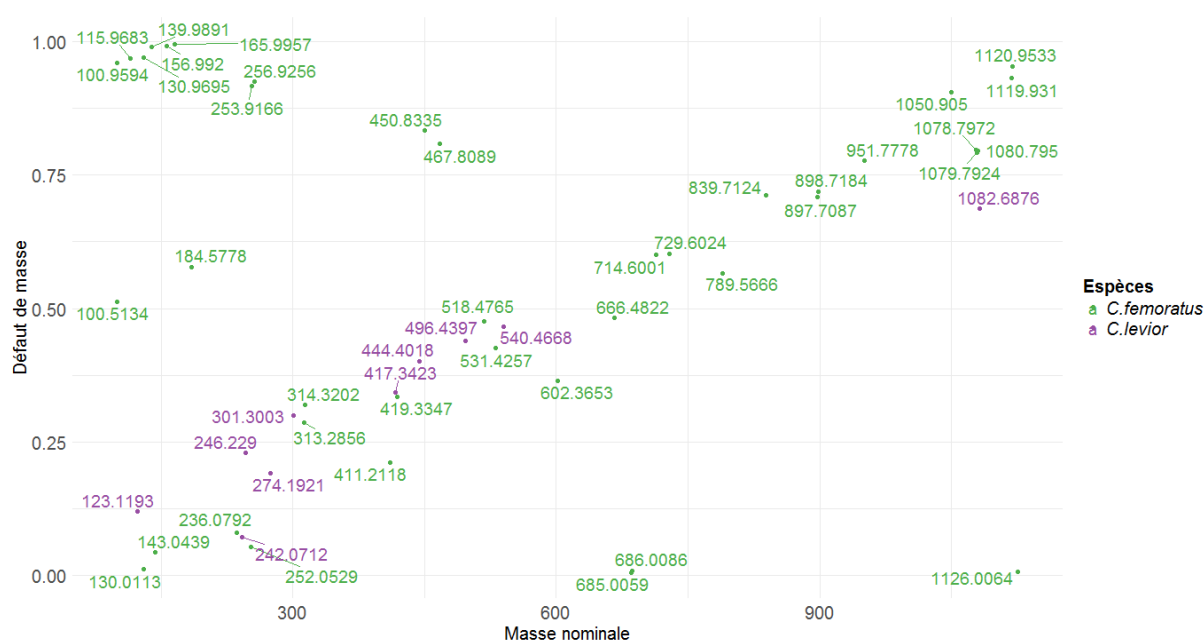


Figure 59 : Analyse de masse de Kendricks des 50 premiers ions sélectionnés par le composant 1 de la sPLS-DA selon l'espèce. Les ions les plus abondants pour une espèce sont de la couleur de l'espèce (*Camponotus femoratus* en vert et *Crematogaster levior* en violet)

Table 11 : Informations chimiques de la base de données Metlin pour 5 ions de *Camponotus femoratus*.

HRMS de l'ion	adduit	delta ppm	formule	famille chimique
898,7184	M+H ⁺	8	C ₅₂ H ₁₀₀ NO ₈ P	PC : phospholipide avec la phosphatidylcholine
839,7124	M+H ⁺	0	C ₅₄ H ₉₄ O ₆	TG : triglycéride
789,5666	M+H ⁺	3	C ₄₃ H ₈₁ O ₁₀ P	PG : phospholipide avec le phosphatidyleglycérol
729,6024	M+H ⁺	0	C ₄₆ H ₈₀ O ₆	TG : triglycéride
602,3653	M+H ⁺	5	C ₂₅ H ₄₇ N ₉ O ₈	peptide (5 acide aminés)

(b) sPLS-DA et site de prélèvement

Le modèle de sPLS-DA avec la variable catégorielle site de prélèvement n'est pas validé car le taux d'erreur de classification est important sur le premier composant (46,8%) et reste conséquent sur le deuxième composant (27,7%). Nous n'avons donc pas effectué d'analyses complémentaires pour identifier les ions discriminants des sites de prélèvement. Ce résultat renforce l'hypothèse selon laquelle les métabolites détectés seraient issus de la fourmi ou de ses potentiels symbiontes transmis verticalement.

E. Futures orientations

En conclusion de ce chapitre, nous avons pu montrer que le microbiote des deux espèces de fourmis semble principalement acquis de l'environnement, mais que certains groupes pourraient être spécifiques à l'espèce voire, constitutifs de la parabiose, bien que des analyses plus fines doivent être menées pour confirmer ce résultat. Malgré cette vision partielle, nos résultats corroborent donc bien avec l'idée que la majorité du microbiote cuticulaire est acquis de l'environnement. Cependant, nos résultats ne nous permettent pas de déterminer si ce microbiote est « bénéfique », i.e. permettrait à l'hôte une protection contre les pathogènes sur un spectre large, car les métabolites obtenus ici semblent être principalement biosynthétisés de l'hôte. Une analyse plus fine en excluant les ions potentiellement issus de l'hôte pourrait apporter de nouveaux éléments. Il est probable que dans ce cas, nous pourrions mieux tester quels sont les métabolites responsables de la faible variance du métabolome expliquée par les sites de prélèvement. Au regard de la forte implication du site de prélèvement dans la composition du microbiote, nous espérons qu'une analyse de corrélation entre OTUs et métabolites permettra de mieux appréhender la fonction effective de ce microbiote fortement contraint par l'environnement. Peu d'études se sont attelées à la tâche d'interpréter conjointement des données de métabarcoding et de métabolomique dans le contexte d'un microbiote. Les données de métabarcoding décrivent la diversité de la communauté bactérienne

à l'aide de petites séquences d'ADN, tandis que la métagénomique permet en plus d'identifier des fonctions spécifiques du microbiote car tout l'ADN du microbiote est séquencé et peut être analysé en lumière des bases de données d'expressions. La métabolomique permettant de connaître l'activité de la communauté à un instant donné, la métabolomique et la métagénomique sont deux outils plébiscités pour être utilisés conjointement, comme le défend l'article « *An Invitation to the Marriage of Metagenomics and Metabolomics* » (Turnbaugh & Gordon 2008).

La suite envisagée pour l'analyse conjointe de ces données est une analyse de réseaux. Les réseaux permettent de représenter les associations pertinentes entre deux matrices (ici microbiote et métabolome) qui sont déduites à l'aide d'un score d'association par paires entre les variables des deux matrices. La visualisation de ce réseau est un graphe dans lequel chaque OTUs ou ions correspond à un nœud lié ou non en fonction de son score d'association avec les autres. Une étude intégrative du métabolome et du microbiome (ensemble du matériel génétique du microbiote) de la surface des muqueuses intestinales chez l'humain a permis de révéler grâce à ce type d'approche des relations significatives métabolite-microbiote. Des relations cohérentes entre des fonctions attribuées au métagénome et les métabolites effectivement détectés suggèrent que la composition du métagénome est prédictive jusqu'à un certain point des métabolites produits par la communautés microbiennes (McHardy *et al.* 2013).

Les auteurs à l'origine du projet mixOmics, utilisé ici, proposent l'utilisation de méthodes multivariées, comme par exemple l'analyse « *sparse* » par la méthode des moindres carrés partiels (s PLS) sur les deux tables de données pour sélectionner des variables pour construire un réseau (González *et al.* 2013). Un autre travail sur la cartographie du métabolome de la peau a également mis en lumière des corrélations entre les métabolites de surface et la diversité bactérienne de la peau. Ainsi à l'aide d'analyses de corrélation spatiale entre ces deux tables, ils ont mis en évidence des taxons bactériens colocalisés avec des métabolites représentés sur un réseau (Bouslimani *et al.* 2015). Ainsi peu d'études à ce jour proposent l'interprétation de données de métabolomique et de diversité bactérienne, mais les quelques études citées ci-dessus mettent l'accent sur l'importance de la visualisation des analyses de corrélations entre deux tables contenant de nombreuses variables. La construction de réseau apparaît être un outil préliminaire essentiel pour l'aide à la visualisation des résultats et à la prise de décision pour formuler des hypothèses de travail.

Par conséquent, les approches citées ci-dessus pour corréler les analyses de métabarcoding et de métabolomique seront testées avec les données pour les deux espèces du jardin de fourmi au cours de l'année 2017

Conclusions générales et perspectives

Ce travail de thèse concerne l'étude du microbiote bactérien cuticulaire des fourmis en Guyane. Afin de mieux comprendre le rôle des bactéries cuticulaires, l'assemblage des communautés bactériennes et la diversité des molécules bactériennes produites, différentes approches ont été utilisées dans ce projet interdisciplinaire.

Le chapitre 1 présente l'isolement, l'identification et la culture d'actinobactéries présentes sur la cuticule de fourmis. Les 43 souches isolées présentent des activités antagonistes *in vitro* contre une souche fongique pathogène pour les insectes (*Beauveria bassiana*) et contre plusieurs souches bactériennes pathogènes pour l'humain. En parallèle, les activités antimicrobiennes des extraits de culture de ces souches ont été évaluées. Notons que la comparaison des deux essais biologiques donne des résultats complémentaires. Ces résultats soulignent l'intérêt d'utiliser différentes approches dans les stratégies de criblage pour identifier des souches candidates productrices de molécules bioactives d'intérêt et soulignent les biais possibles dans la sélection de souches candidates selon le test biologique mis en place.

Le fractionnement bioguidé effectué à partir de l'extrait de culture d'une de ces souches d'actinobactérie a permis l'isolement et l'identification structurale d'un dipeptide cyclique (Cyclo(Pro-Phe)) possédant des activités antibactériennes documentées. Ces résultats sont particulièrement limités par le temps nécessaire pour évaluer les activités pharmacologiques, pour effectuer le travail de déréplication et pour réaliser les nombreuses analyses chimiques nécessaires à la détermination structurale des composés. Par contre il apparaît que l'analyse des réseaux moléculaires par spectrométrie MS/MS représente une approche rapide et efficace pour évaluer la diversité des composés chimiques connus ou originaux présents dans un extrait de culture bactérienne en limitant considérablement le travail de déréplication chronophage tout en guidant l'analyse ciblée des métabolites d'intérêt. Cette méthode est dépendante des données spectrales présentes dans les bases de données, cependant l'augmentation constante du nombre molécules introduites dans ces bases de données renforce la pertinence de cette approche (Wang *et al.* 2016).

Nous avons également montré l'importance de considérer la coculture des micro-organismes isolés dans un même environnement, ici la cuticule des fourmis, pour induire l'expression des voies de biosynthèses microbiennes restées silencieuse dans les monocultures. L'utilisation des réseaux moléculaires générés par les spectres de fragmentation des composés en spectrométrie de masse a permis d'identifier rapidement les composés produits uniquement dans les cocultures bactériennes. Ces résultats suggèrent que les approches couramment

utilisées actuellement en chimie des substances naturelles atteignent une limite dans l'exploration moléculaire correspondant au plafond observé dans l'identification de composés nouveaux (Newman & Cragg 2012). Ce travail est précurseur et représente un premier exemple de l'utilisation des réseaux moléculaires pour l'étude de cocultures microbiennes. Au-delà des actinobactéries cuticulaires de fourmis, l'intégration de ce type d'approche est primordiale dans le domaine de la chimie des substances naturelles pour assurer une meilleure évaluation de la diversité chimique microbienne.

Dans le deuxième et le troisième chapitre, nous sommes intéressés à la composition des communautés bactériennes cuticulaires dans l'objectif de mieux appréhender le microbiote potentiellement bénéfique de la cuticule des fourmis. En effet, les interactions entre le monde microbien et le monde animal sont de plus en plus explorées (McFall-Ngai *et al.* 2013) et certains suggèrent que ces deux « mondes biologiques » sont indissociables (Rosenberg & Zilber-Rosenberg 2016). Il est généralement admis que les approches interdisciplinaires participent grandement à une meilleure compréhension des communautés microbiennes et des interactions hôtes-microbiote (Franzosa *et al.* 2015). Ainsi, c'est avec un grand intérêt, après avoir exploré la mise en culture d'actinobactéries et la diversité chimique de leurs extraits, que nous avons décidé de réaliser l'analyse par métabarcoding du microbiote bactérien cuticulaire des fourmis de Guyane.

Tout comme les sciences forensiques ou dédiées à l'archéologie, l'étude du microbiote cuticulaire d'insectes de petite taille nécessite de mettre au point des protocoles de biologie moléculaire adaptés à des échantillons contenant peu de matériel génétique et particulièrement sensibles aux contaminations. Après de nombreux échecs d'extraction d'ADN bactérien de la cuticule des fourmis, nous avons décidé de réaliser une étude méthodologique visant à comparer quatre méthodes d'extraction d'ADN différentes à partir de deux espèces de fourmis *Atta cephalotes* et *Pseudomyrmex penetrator* de taille différente. Dans un premier temps nous avons évalué la richesse et la composition du microbiote cuticulaire ainsi que les biais et les contaminations liés aux quatre méthodes d'extraction. Les résultats montrent que ces quatre méthodes donnent des résultats satisfaisants et similaires pour les fourmis *Atta cephalotes* de grande taille tandis que de plus grandes variabilités et reproductibilités de composition des communautés sont observées pour les fourmis de petite taille *Pseudomyrmex penetrator* selon la méthode utilisée. Cette étude montre que l'utilisation du kit QIAamp, et dans une moindre mesure MoBio, permettent d'accéder à la composition du microbiote bactérien ainsi qu'aux variabilités interindividuelles pour les deux espèces. En comparant les communautés bactériennes d'individus uniques et d'individus regroupés en pool nous n'avons obtenu de

différences significatives dans la composition des communautés bactériennes pour les deux espèces étudiées. A l'inverse la comparaison entre des individus de petites et de grandes tailles chez les fourmis *Atta cephalotes*, nous avons mis en évidence des différences des communautés bactériennes cuticulaires. Ce dernier résultat est très intéressant car aucune littérature ne mentionne de variabilité du microbiote cuticulaire bactérien selon les castes chez les fourmis attines. Ainsi, il serait pertinent de réaliser une étude du microbiote cuticulaire de l'ensemble des castes d'une colonie d'*Atta cephalotes* afin de mieux comprendre l'assemblage et la transmission des communautés bactériennes cuticulaires au sein de cette espèce. Nos résultats soulignent l'importance des contrôles d'extraction, du nettoyage des données et d'intégrer des individus homogènes au sein d'une classe définie. Ces travaux ouvrent de nouvelles perspectives dans l'étude du microbiote cuticulaire chez les insectes en général.

Les insectes constituent d'excellents modèles d'études pour adresser les questions hôte-microbiote et accéder aux fonctions de ces interactions, notamment bénéfique. Bien que la plupart de ces études s'intéressent au microbiote de l'insecte (extraction à partir de l'insecte entier) ou de son système digestif (dissection des différents intestins), le protocole que nous proposons pour la cuticule des fourmis pourrait être utilisé par la communauté des entomologistes à partir des cuticules d'autres espèces d'insectes.

En utilisant le protocole décrit précédemment nous avons étudié dans le troisième chapitre la composition du microbiote bactérien cuticulaire des deux espèces de fourmis *Camponotus femoratus* et *Crematogaster levior* habitant dans les jardins de fourmis et partageant le même nid localisé dans le réseau racinaire de plantes disposées par les fourmis dans un arbre. L'étude de cette parabiose est pertinente car elle permet de tester la transmission verticale et la variabilité du pool bactérien environnemental. En effet la variabilité de la composition des communautés bactériennes liée à l'espèce mise en évidence chez *Atta cephalotes* et *Pseudomyrmex penetrator* dans le chapitre 2 est confondante avec celle liée à leur habitat. Dans le cas de *Camponotus femoratus* et *Crematogaster levior*, l'échantillonnage mené cette fois sur trois sites distincts en Guyane et l'effet habitat est contrôlé puisque ces deux espèces de fourmis partagent le même nid. Les analyses de métabarcoding mettent en évidence que le site de prélèvement des individus est le facteur principal expliquant la variance de la composition des communautés bactériennes cuticulaires. Ce résultat suggère qu'une majeure partie du microbiote bactérien cuticulaire de ces deux espèces est acquis dans l'environnement. Néanmoins, une plus faible partie de cette variance est expliquée par le facteur espèce. Ce résultat suggère que l'hypothèse d'une transmission verticale d'une partie du microbiote bactérien cuticulaire ne peut pas être rejetée.

En parallèle, avec pour objectif mettre en lumière un rôle bénéfique du microbiote cuticulaire en détectant la présence de molécules microbiennes bioactives, nous avons réalisé l'analyse métabolomique des cuticules des individus des deux espèces habitant dans les jardins de fourmi. Bien que l'identification des métabolites par spectrométrie de masse haute résolution non combinées avec des analyses MS/MS soit délicate par simple comparaison des masses des ions avec les bases données, nous avons obtenu des résultats intéressants. Premièrement, il s'avère que l'espèce de fourmi à laquelle appartiennent les individus explique les différences de composition des métabolites entre les individus de *Camponotus femoratus* et *Crematogaster levior*. Les tentatives d'identification des métabolites les plus spécifiques à chaque espèce furent limitées. Cependant plusieurs molécules de la famille des phospholipides identifiées grâce aux bases de données semblent spécifiques à *Camponotus femoratus*. Ces phospholipides ne sont vraisemblablement pas d'origine bactérienne. Ils sont synthétisés *de novo* par la fourmi et semblent plutôt associées aux mécanismes de transport des CHCs des cellules œnocytes jusqu'à la surface de la cuticule. En complément des analyses chimio-taxonomiques des CHCs par GC/MS, l'analyse par UPLC/ESI-MS montre pour la première fois que les phospholipides cuticulaires pourraient être spécifique au genre ou à l'espèce de fourmi étudiée. Deuxièmement, les résultats montrent qu'une partie des différences de composition des métabolites des individus est expliquée par le site de prélèvement. Puisque d'après nos analyses le microbiote bactérien des deux espèces apparaît être significativement acquis dans l'environnement, ces données métabarcoding et métabolomique permettent de suggérer que la production de métabolites d'origine bactérienne est possible sur la cuticule de ces fourmis.

L'assemblage des deux tables de données métabolomique et métabarcoding n'a pas pu être initié dans travail de thèse notamment pour des questions de temps et par la nécessité de mener une réflexion sur le choix des analyses statistiques. Néanmoins, ce travail offre des perspectives et représente le premier exemple d'étude combinée du microbiote cuticulaire de fourmis. Il est donc important de poursuivre ce travail qui tend à montrer la présence d'un microbiote bénéfique via la détection *in vivo* des métabolites microbiens de surface. Nous discutons ici des méthodologies possibles dans ce domaine.

D'autres « omics » pourraient être mises en place pour révéler le rôle bénéfique du microbiote cuticulaire des fourmis. Ainsi, une approche consisterait à réaliser une analyse métabarcoding des cuticules à partir de plusieurs espèces de fourmis phylogénétiquement différentes. Ceci permettrait de mieux comprendre les facteurs impactant l'assemblage des communautés bactériennes cuticulaires, par exemple le régime alimentaire, le type d'habitat ou de degré de sociabilité. Dans ce contexte lors des dernières collectes réalisées en Guyane dans

le cadre ce travail de thèse, sept espèces de fourmis ont été collectées (Table 12). L'objectif est de comparer la composition des communautés bactériennes entre des espèces phylogénétiquement proches et vivant dans un environnement différent et des espèces phylogénétiquement différentes et vivant dans un environnement similaire. Dans les résultats du chapitre 3, il semblerait que pour les jardins de fourmis la composition des communautés bactériennes soit principalement contrainte par l'environnement. Les données de métabolomique ainsi que les données de métabarcoding pour les sept autres espèces ont été acquises respectivement lors des dernières injections réalisées à Paris en août 2016 et lors du dernier séquençage haut débit réalisé à Toulouse en juin 2016. L'interprétation des données sera réalisée au cours de l'année 2017.

Table 12 : Caractéristiques des 9 espèces de fourmis pour l'étude à plus grande échelle.

catégories	espèces	milieu	régime alimentaire	glandes	
				venin	métapleurale
espèces proches/environnement différent	<i>Ectatomma brunneum</i>	sable	généraliste	présente	présente
	<i>Ectatomma brunneum</i>	terre	généraliste	présente	présente
	<i>Pseudomyrmex penetrator</i>	myrmécophyte	herbivore/généraliste	présente	présente
	<i>Pseudomyrmex termitarius</i>	terre	généraliste	présente	présente
	<i>Atta sexdens</i>	terre	herbivore	?	présente
	<i>Atta cephalotes</i>	terre	herbivore	?	présente
espèces différentes/environnement similaire	<i>Crematogaster levior</i>	jardin de fourmi	herbivore/généraliste	présente	présente
	<i>Camponotus femoratus</i>	jardin de fourmi	généraliste	présente	absente
	<i>Paraponera clavata</i>	terre	généraliste/prédateur	présente	présente
	<i>Gigantiops destructor</i>	terre	généraliste	?	présente

Ensuite il serait intéressant de réaliser une étude métagénomique pour mieux comprendre les fonctions bénéfiques de ces bactéries cuticulaires associées à la production de molécules bioactives d'intérêt. En reconstituant le génome des bactéries et en utilisant les bases de données disponibles, il est possible d'identifier des clusters de gènes codant pour des métabolites connus pour leurs activités antibiotiques (Reddy *et al.* 2014). Par exemple, les gènes codant pour la biosynthèse des peptides non ribosomiques ou des polycétides pourraient être ciblés. En effet ces familles chimiques ont donné lieu aux molécules thérapeutiques prescrites les plus actives contre les bactéries pathogènes chez l'humain résistantes aux antibiotiques. La découverte de nouvelles molécules au sein ces familles chimiques est l'une des stratégies principales dans la lutte contre les bactéries pathogènes (Charlop-Powers *et al.* 2016). Sachant que de nombreux gènes du génome bactérien restent silencieux dans les approches de culture bactérienne, la métagénomique est une stratégie prometteuse pour accéder à l'intégralité du matériel génétique. Néanmoins, bien qu'une grande part des gènes de biosynthèse ne soient pas identifiés dans les bases de données, ceux qui le seront peuvent être intégrés dans le génome de bactéries ou levures utilisés pour l'expression hétérologue des gènes. Il est donc alors possible de connaître les modifications chimiques originales dans les structures chimiques des peptides

non ribosomiques ou des polycétides. Bien que la métagénomique ne permette pas de savoir si ces clusters de gènes sont réellement exprimés, l'étude transcriptomique des cuticules pourrait s'avérer pertinente. Au-delà de l'intérêt pharmacologique, il serait intéressant d'un point de vue écologique de savoir la proportion de gènes effectivement exprimée et ainsi caractériser plus profondément le potentiel rôle bénéfique du microbiote bactérien des cuticules de fourmis. Enfin, il est important de poursuivre les analyses métabolomiques des cuticules de fourmis pour identifier le métabolome bactérien produit *in vivo*. Cependant afin d'identifier un grand nombre de molécules il sera nécessaire d'avoir recours aux réseaux moléculaires par MS/MS et de développer des approches pour corréler la présence de ces métabolites avec les analyses génétiques des génomes bactériens cuticulaires.

Annexes

Annexe 1 : Protocole du test cross-streak.....	157
Annexe 2 : Protocole de préparation des extraits.	159
Annexe 3 : Protocole des tests antibiotiques pourcentage d'inhibition des extraits à 400 µg/mL.....	161
Annexe 4 : Pourcentage d'inhibition des extraits à 400 µg/mL.	162
Annexe 5 : Spectres RMN de la dicétopipérazine ((R, S)Pro-(R, S)Phe).	163
Annexe 6 : Supporting Information de l'article Comparative analysis of DNA extraction methods to study the body surface microbiota of insects: a case study with ant cuticular bacteria	168
Annexe 7 : Table des OTUs communs aux deux espèces de fourmis (remplissage orange) et aux deux espèces de fourmis et au nid (remplissage jaune) pour les trois sites....	180

Annexe 1 : Protocole du test cross-streak

Les souches d'actinobactéries sont inoculées sous la forme d'une bande centrale sur du milieu YMEA solide à partir de pré-cultures en milieu YMEA liquide. Elles sont incubées à 30°C pendant 7 jours : ce temps est nécessaire pour que les souches d'actinobactéries puissent produire les métabolites qui se diffuseront dans le milieu. La pré-culture liquide est une culture de 24h de la souche d'actinobactérie prélevée de son milieu YMEA agar de stockage dans 15 mL de milieu liquide YMEA à 30°C sous agitation. La bande centrale a une longueur de 8.5 cm, une largeur de 1 cm et est réalisée avec 32 µL de pré-culture étalée à l'aide d'un ose stérile. Afin de dessiner des bandes d'actinobactéries de manière homogène entre toutes les boîtes, une feuille de papier millimétré avec les marquages est attachée en dessous de la boîte de pétri pour faciliter le traçage et par la suite la lecture des résultats. Chacune des boîtes de monocultures sont réalisées en duplicate.

Le test du cross-streak consiste à inoculer perpendiculairement à la bande d'actinobactérie des bandes de « souches tests » dont on observera la croissance par rapport à celle d'un témoin négatif (sans l'actinobactérie). Les « souches tests » sont les souches pathogènes pour l'homme ou la fourmi (entomopathogène) à partir desquelles l'activité d'inhibition des actinobactéries est observée. Dans le cadre de ces tests, 5 pathogènes humains et une souche fongique entomopathogène sont sélectionnés.

Milieu Luria Bertani (LB) : 10 g de tryptone (pancreatic digest of casein), 5 g de yeast extract, 5 g de NaCl, 1 L d'eau Milli Q. Milieu Potatoes Glucose Agar (PDA) : 4 g de potato extract, 20 g de dextrose, 1 L d'eau Milli Q. Milieu Sabouraud : 10 g de peptone pepsique de viande, 35 g de glucose, 1 L d'eau Milli Q.

Pour les bactéries et la levure : une pré-culture de la souche test est préparée la veille en ajoutant un ose de cette souche dans 4 mL de son milieu de référence à 30 °C. Des cultures liquides de concentration équivalente entre tous les tests d'une même souche sont nécessaires. Ainsi la pré-culture est diluée au 1/100^{ème} pour suivre la croissance microbienne à l'aide de la lecture de la DO (densité optique) à 560 nm contre un témoin de milieu. Lorsque la croissance atteint une DO de 0,2 (phase exponentielle), la culture liquide est diluée au 1000^{ème} et directement appliquée perpendiculairement à l'actinobactérie à l'aide d'un ose. Les boîtes sont incubées à 30°C pour les bactéries et 25 °C pour la levure.

La distance entre les bandes des souches tests et la bande d'actinobactérie est de 0.6 cm, la bande des souches tests est de 3 cm de longueur, de 0.5 cm de largeur et la distance entre les bandes de deux souches tests est de 1.5 cm. Un maximum de 4 souches tests est testé par plaque.

Un contrôle négatif est réalisé de la même manière avec les souches tests sans l'actinobactérie. Afin de dessiner des bandes de souche tests de manière homogène entre toutes les boîtes, une feuille de papier millimétré avec les marquages est attachée en dessous de la boîte de pétri pour faciliter le traçage et par la suite la lecture des résultats.

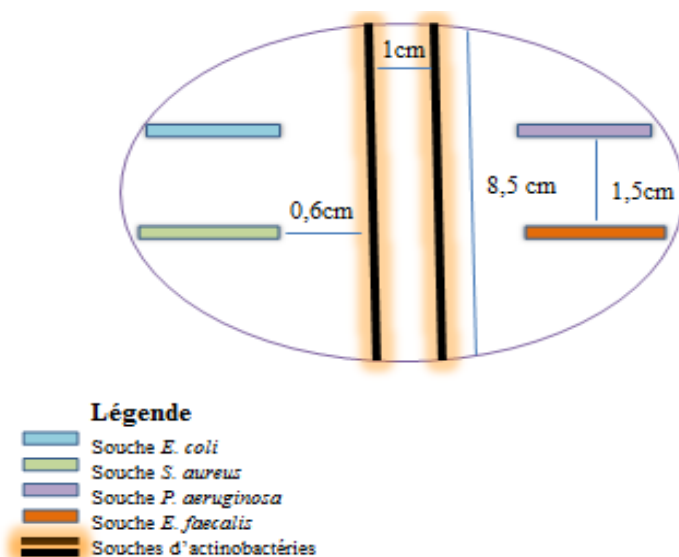


Figure 60 : Schéma de la méthode cross-streak.

Pour le champignon : les spores sont prélevées à partir d'une culture vieille de 15 jours du champignon en ajoutant 3 mL de sérum physiologique sur la boîte de pétri contenant le champignon. Un öse stérile est utilisé pour froter le champignon et décrocher les spores. 1 mL de la suspension de spores est prélevé puis dilué au 1/10^{ème}. Les spores sont comptées au microscope optique (grossissement *40) avec la cellule de comptage de Neubauer. La concentration en spore est ajustée à 10⁶ spores/mL avec du sérum physiologique. Puis 15 µL de cette suspension de spores sont étalés perpendiculairement à l'actinobactérie, les étapes suivantes sont identiques à celles expliquées ci-dessus pour les bactéries et les levures. Les boîtes sont incubées à 25 °C.

La lecture des résultats est faite à 24h, 48h et 72h pour les bactéries et uniquement à 48h pour la levure et le champignon. A l'aide d'un papier calque millimétré (dessiné à partir du modèle ayant servi à la mise en culture des souches tests) le nombre de carré de 1 mm de côté remplissant la zone d'inhibition est compté. Une bande de souche test non inhibée, par exemple celles présentes sur le témoin négatif, recouvre une surface de 150 mm². Le pourcentage d'inhibition d'une actinobactérie pour chaque souche testée est obtenu avec cette formule :
pourcentage d'inhibition = (zone inhibée) / (zone totale) * 100.

Annexe 2 : Protocole de préparation des extraits.

Production des métabolites (monoculture)

- 1) pré-culture : la souche pure est introduite dans un erlenmeyer de 250 mL contenant 50mL de milieu YMEA liquide stérile, sous agitation pendant 48 heures à 30°C.
- 2) cultures : 2 mL de la pré-culture sont introduits dans un erlenmeyer de 500mL contenant 150 mL de milieu YMEA liquide stérile. 2 réplicats sont réalisés de manière identique. Les cultures sont agitées pendant cinq jours à 30°C.
- 3) arrêt de la prolifération bactérienne : les cultures sont transvasées dans des ballons en verre de 500mL rapidement plongés dans un bain d'azote liquide afin de stopper le métabolisme et d'éviter les dégradations enzymatiques.
- 4) lyophilisation : le mélange congelé contenant le milieu de culture avec les débris de cellules bactériennes et les métabolites produits est lyophilisé.

Production des métabolites (cocultures) :

Plusieurs souches d'actinobactéries isolées à partir d'une même espèce de fourmi sont mises en culture deux à deux, à trois ou plus.

- 1) pré-cultures : la mise en culture en milieu liquide YMEA se fait de manière individuelle pour chaque souche. La souche pure est introduite dans un erlenmeyer de 250 mL contenant 50 mL de milieu YMEA stérile, sous agitation pendant 48 heures à 30°C.
- 2) cocultures : 2 mL de chaque pré-culture choisies sont introduits dans un erlenmeyer de 500 mL contenant 150 mL de milieu YMEA liquide stérile. 2 duplicatas sont réalisés de manière identique. Les cocultures sont agitées pendant cinq jours à 30°C.
- 3) l'arrêt de la prolifération bactérienne et la lyophilisation se font de la même manière que pour les monocultures.

Extraction des métabolites et purification SPE :

Un extrait est réalisé de manière identique pour chaque réplica de culture, c'est-à-dire que l'on obtient 3 extraits différents issus des 3 réplicats de cultures (mono ou cocultures). La variabilité de la production de métabolites par les bactéries étant importante, il est absolument nécessaire de réaliser ces répétitions.

Extraction des métabolites : 600 mg de lyophilisat sont extraits aux ultra-sons pendant 20 min dans 20 mL d'un mélange de solvant **dichlorométhane/méthanol/eau** (64/36/8) dans un erlenmeyer en verre de 50 mL. L'extrait est filtré sur coton puis le filtrat est séché avec 120 mg de C₁₈ à l'évaporateur rotatif.

Purification de l'extrait : une colonne de SPE C18 de 500 mg est conditionnée à l'eau. L'extrait brut est ajouté en haut de la colonne sous forme d'un dépôt sec. Le premier lavage consiste à passer 6 mL d'eau en deux fois afin d'éliminer un maximum de sucres du milieu de culture. Les métabolites sont élués avec 3 mL de méthanol et séchés au concentrateur Speed Vac dans un tube en verre. Les extraits purifiés sont conservés à -20°C.

Préparation des extraits purifiés pour l'analyse chimique en HPLC/HRMS : solution des extraits purifiés à 1 mg/mL dans du méthanol filtrées à 0.20 µm conservés à -20°C.

Les tests antibiotiques : des solutions mères des extraits purifiés sont préparées dans du DMSO à 20 mg/mL et conservées à -20°C.

Annexe 3 : Protocole des tests antibiotiques pourcentage d'inhibition des extraits à 400 µg/mL

Quatre souches bactériennes pathogènes pour l'homme sont sélectionnées et cultivées en milieu liquide LB : *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus* et *Enterococcus faecalis*. Chaque extrait purifié est testé de manière identique dans 3 puits à 400 µg/ml. Les tests antibactériens sont réalisés dans des plaques 96 puits suivant un plan de plaque optimisé (Wiegand *et al.* 2008).

- Suspension bactérienne : les tests débutent avec des suspensions bactériennes diluées au 1/200 après la lecture de la densité optique (DO) à 0,2 UA pour toutes les souches pour des raisons de reproductibilité. 90 µL de cette suspension bactérienne sont ajoutés à tous les puits des extraits à tester ainsi que des témoins négatifs et positifs.
- Les extraits purifiés sont en solution à 20 mg/mL dans du DMSO. Dans un puits, 9 µL de cette solution sont mélangés à 45 µL de LB (dilution 1/5) = solution fille où il y a 20% de DMSO et où la concentration est de 4 mg/mL. Puis 10 µL de cette solution fille sont ajoutés dans 3 puits pour chaque extrait à tester = l'extrait est testé à la concentration de 400 µg/mL et le DMSO est présent à 2% une fois les 90 µL de suspension bactérienne ajoutée (dilution au 1/10).
- Trois types de témoins sont ajoutés : un témoin de milieu LB pour s'assurer de la stérilité du milieu de culture, un témoin positif contenant du milieu LB et du DMSO à 2% et un témoin négatif contenant du Sodium Dodecyl Sulfate (SDS) à 5% perturbant la paroi bactérienne et conduisant à la lyse cellulaire.
- La plaque est incubée sous agitation pendant 24 heures à 30°C, puis la lecture des DO de la plaque se fait avec un lecteur de plaque TriStar LB 941 BERTHOLD utilisant le logiciel MikroWin 2000 à une longueur d'onde de 620 nm.

	1	2	3	4	5	6	7	8	9	10	11	12
A	eau	Xf	Yf							T-	T+	T LB
B	eau	X1	Y1							T-	T+	T LB
C	eau	X2	Y2							T-	T+	T LB
D	eau	X3	Y3							T-	T+	T LB
E	eau	Z3								T-	T+	T LB
F	eau	Z2								T-	T+	T LB
G	eau	Z1								T-	T+	T LB
H	eau	Zf								eau	eau	eau

Figure 61 : Schéma d'une plaque de 96 puits de test antimicrobien.

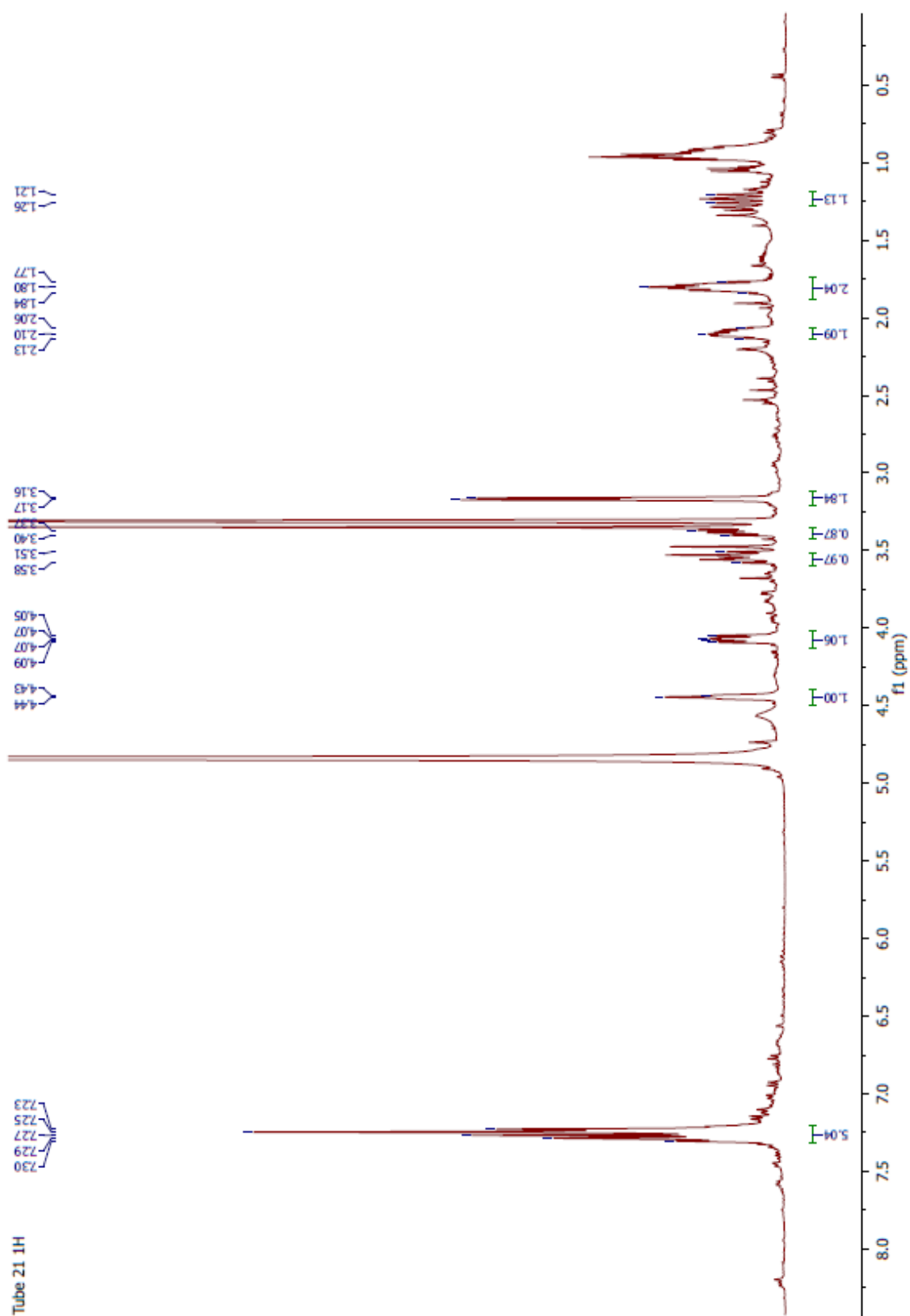
Les résultats de la lecture des plaques permettent de calculer le pourcentage d'inhibition de l'extrait à 400 µg/mL pour les bactéries pathogènes considérées.

Annexe 4 : Pourcentage d'inhibition des extraits à 400 µg/mL.

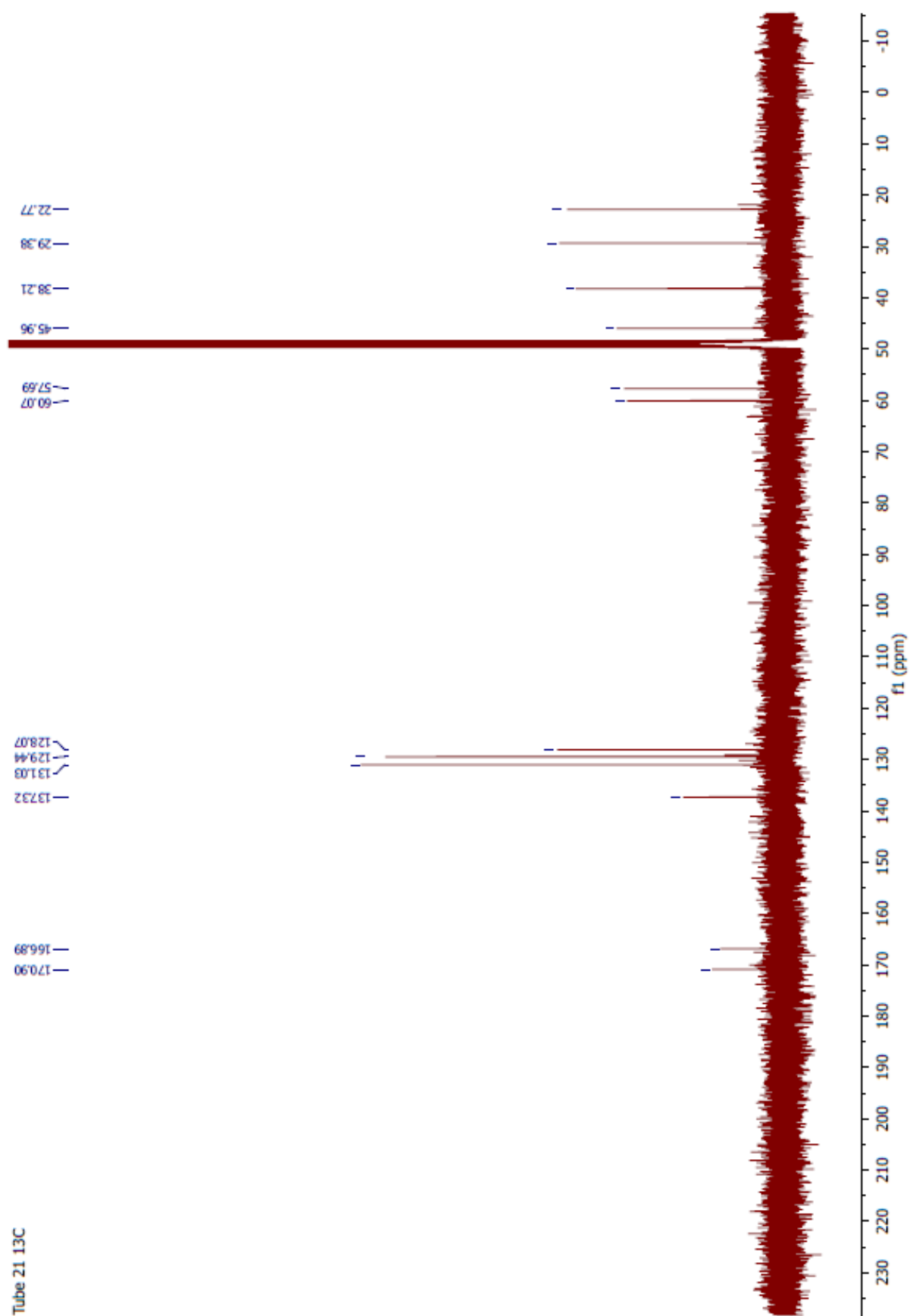
Codes	Pourcentage d'inhibition des extraits à 400 µg/ml											
	<i>Escherichia coli</i>			<i>Pseudomonas aeruginosa</i>			<i>Staphylococcus aureus</i>			<i>Enterococcus faecalis</i>		
	1	2	3	1	2	3	1	2	3	1	2	3
32-41	16,4 ± 0,2	15,5 ± 2,6	20,1 ± 0,9									
32-51												
32-62												
34-61												
38-11												
38-22												
40-11								26,5 ± 0,8	27,6 ± 1,7			
40-62												
41-52							92,3 ± 1,2	97,7 ± 1,2	100,0		90,0 ± 0,8	100,0
43-62							100,0	93,3 ± 1,4	73,7 ± 1,1	100,0	48,8 ± 1,7	61,5 ± 0,1
44-313												
45-21												
46-41										42,1 ± 2,5		18,5 ± 1,6
46-51	43,2 ± 0,2	32,6 ± 1,8	41,9 ± 1,3									
48-13												
48-14								24,5 ± 2,6			37,3 ± 0,9	30,8 ± 1,6
48-212										42,9 ± 2,2		
48-31										28,9 ± 5	15,4 ± 0,2	14,2 ± 2,3
48-43												
51-111												
51-62												
53-12												
53-41								21,4 ± 1,6			14,9 ± 0,1	50,7 ± 0,4
59-11												
62-41												
444												
447												
7811												
7819	23,5 ± 1,0	25,1 ± 0,9										
7821												
518												
492							80,3 ± 1,2	32,5 ± 0,3			36,2 ± 1,2	60,3 ± 0,2
497							98,7 ± 1,8		82,3 ± 2,2			59,1 ± 0,6
498												
4916												
555												
5510										33,5 ± 0,3		34,8 ± 2,8
5513							100,0	94,2 ± 3,5		79,4 ± 0,9	98,0 ± 1,6	63,6 ± 1,9
564												
566												
581												
788												
902												
9318												
2914											22,1 ± 0,8	61,3 ± 0,2
417							100,0		100,0	100,0	100,0	100,0
CoC1	83,8 ± 9,9		41,1 ± 0,9	14,5 ± 1,3	17,3 ± 1,3		100,0	80,3 ± 5,1	100,0	91,6 ± 0,2	25,9 ± 0,6	95,1 ± 0,4
CoC2								95,9 ± 0,4		51,3 ± 0,6	83,8 ± 0,4	
CoC3	19,8 ± 1,7						100,0	100,0	100,0	93,4		
CoC4							100,0	100,0		45,1 ± 1,3	77,9 ± 0,5	
CoC5												
CoC6									87,4 ± 0,4	72,1 ± 0,8		64,8 ± 0,4

Annexe 5 : Spectres RMN de la dicétopipérazine ((R, S)Pro-(R, S)Phe).

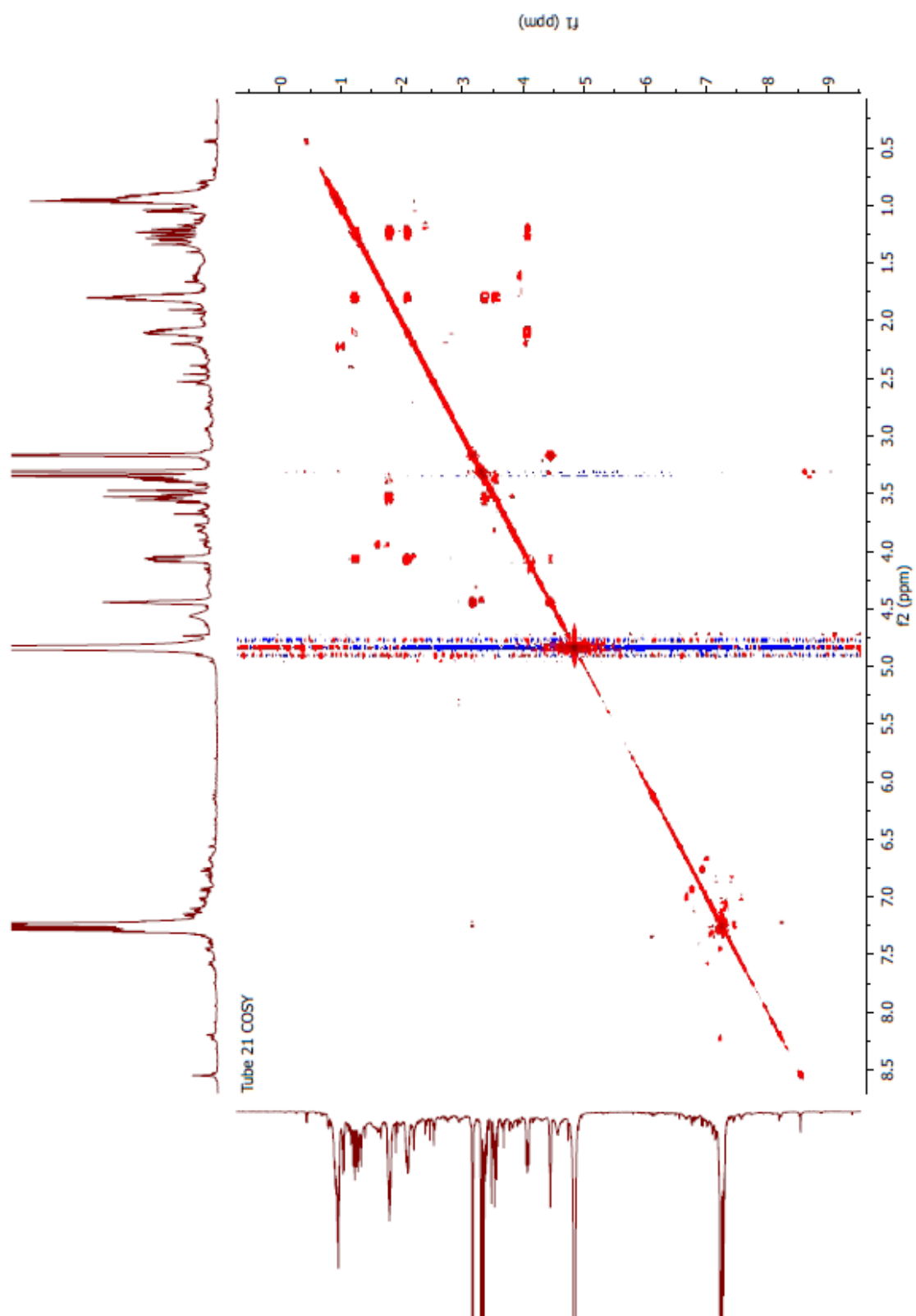
(1) ^1H



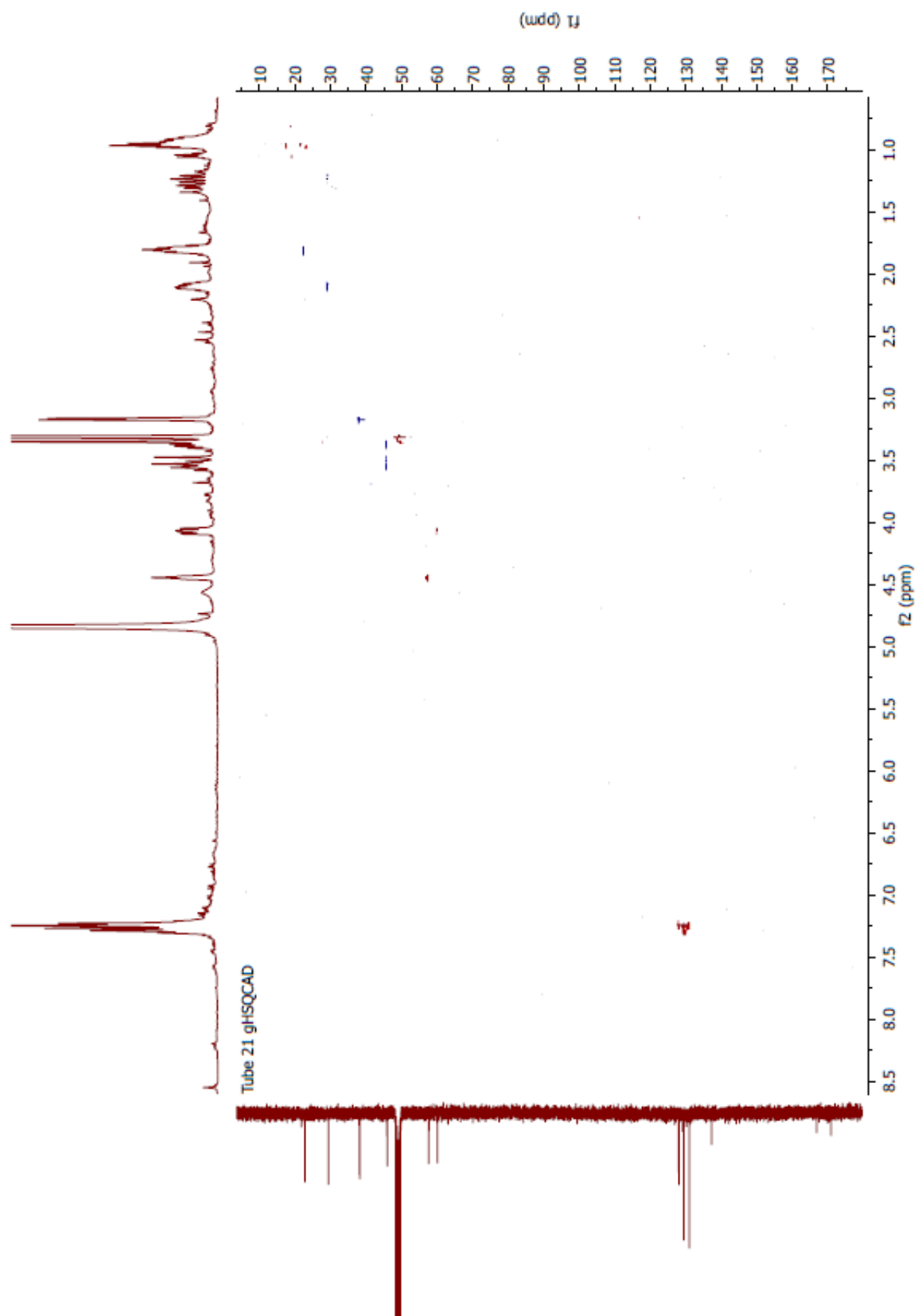
(2) 1C



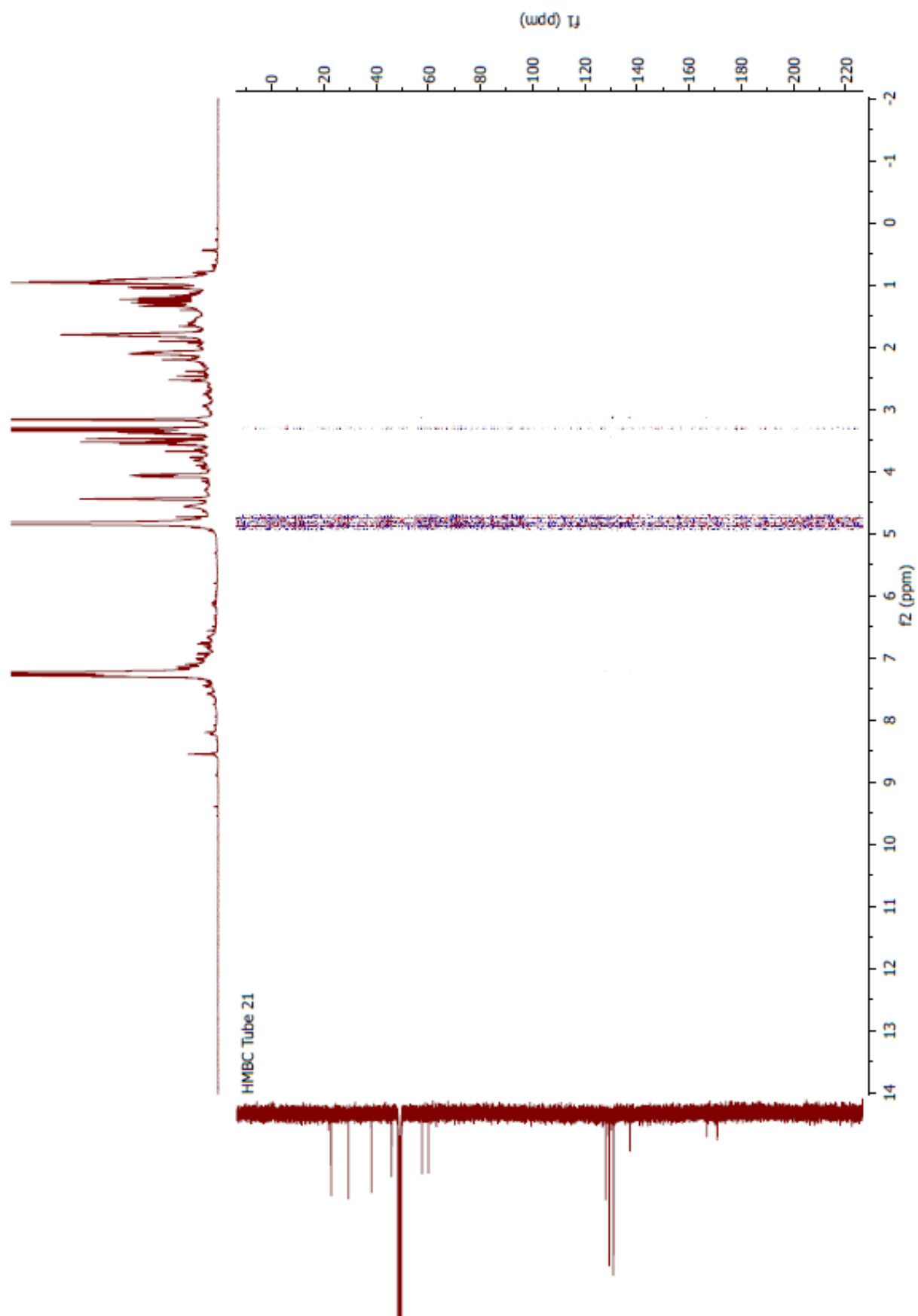
(3) gCOSY



(4) gHSQCAD



(5) gHMBC



Annexe 6 : Supporting Information de l'article Comparative analysis of DNA extraction methods to study the body surface microbiota of insects: a case study with ant cuticular bacteria

Appendix S1. Ant washing and DNA extractions protocols

The MO BIO extraction method differs from the other methods in that it contains its own lysis buffer. Ants were directly vortexed for 1 min in 600 µL of the lysis MO BIO buffer, and the DNA extraction method followed manufacturer instructions with minor modifications to improve the cell wall rupture: a step was added where ant MO BIO lysis buffer mixes were heated to 65 °C for 10 min, as recommended by Fierer et al. (Fierer *et al.* 2008).

For the other three DNA extraction methods, QIAamp, Dneasy, and PC, bacterial cells were dislodged by washing ants in a 1.5 mL plastics tubes containing 300 µL of Buffer A (20 mM Tris·Cl, pH 8.0, 2 mM sodium EDTA, 1.2% Triton® X-100) and vortexed for 1 min. Triton X-100 is a surfactant that helps to detach cells from the cuticle. For pool samples, 10 individuals were washed in 300 µL of Buffer A. Ants were removed with sterilized tweezers from the tubes and the bacterial suspension was kept at -20 °C until further processing.

The cell lysis in QIAamp, Dneasy, and PC methods consists of adding 15 µL of lysozyme (Carl Roth, Karlsruhe, Germany) at 0.5 g/mL to the bacterial suspension, and incubating the solution for 1h at 37 °C, in accordance with the modification recommended by the manufacturer. Indeed, lysozyme is particularly useful for the lysis of Gram+ bacteria. Then 200 µL of Qiagen buffer AL and 25 µL of proteinase K, both provided by the Qiagen kits (QIAamp, Dneasy), were added. These reagents cleave peptide bonds of proteins surrounding the DNA and inactivate nucleases. After gentle mixing, the samples were incubated for 1h at 56 °C.

The DNA isolations in QIAamp and Dneasy methods followed manufacturer instructions in which DNA is bound to silica-membranes, washed, and eluted. QIAamp and Dneasy kits differ by the pore size of the silica-membrane. QIAamp kit has smaller pore size, which limits DNA losses

For PC method, 300 µL of phenol and 100 µL of chloroform were added to the lysate solution. After centrifugation at 1400 rpm for 10 min, the aqueous phase is washed with 300 µL of chloroform to remove any residual phenol. Cold sodium acetate (3 M; pH 5.2; 1/10 of the volume of aqueous phase) was added to neutralize the DNA. DNA was precipitated in ethanol (2/3 vol. absolute ethanol at -20 °C

overnight). After centrifugation at 1400 rpm for 10 min, the solvent phase was carefully removed, and 500 μ L of 70% ethanol was added to rehydrate DNA. Samples were then centrifuged at 1400 rpm for 10 min and the supernatant was carefully removed. DNA pellets were suspended in 20 μ L of UHQ water and stored at -20°C for downstream analyses.

Figure S1. Effect of DNA extraction methods on the number of reads (a) and OTUs (b) in *A. cephalotes* (blue), *P. penetrator* (yellow) with raw data. The y axis is in a logarithmic scale. The figure includes extraction control samples (dark grey) for comparative purposes.

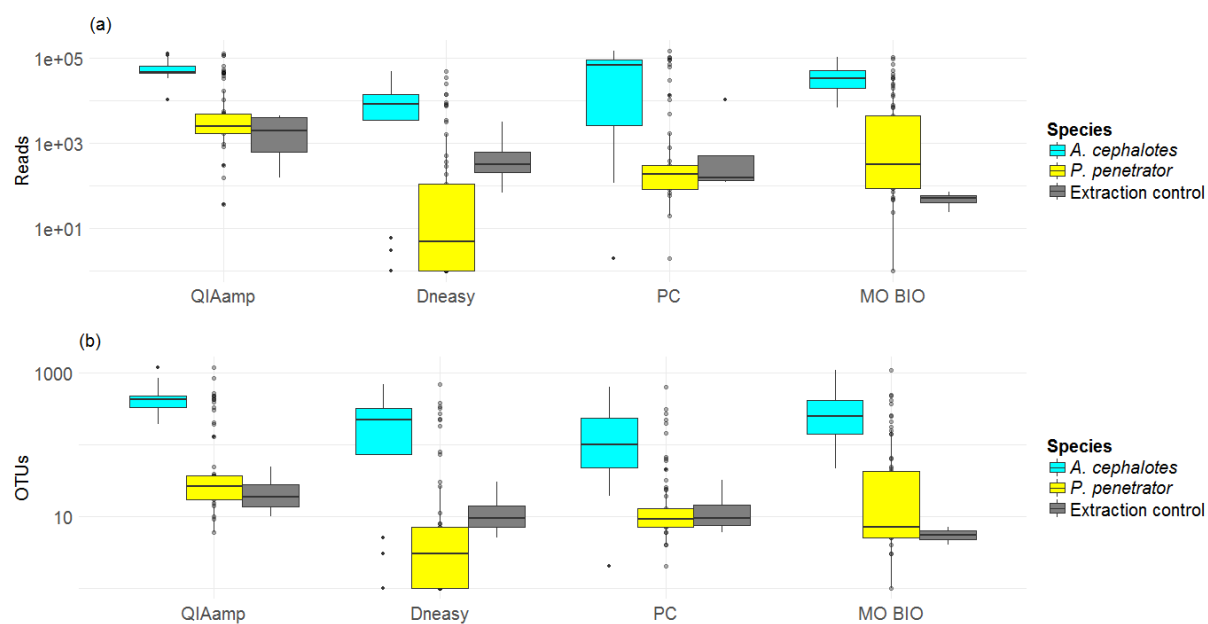


Figure S2. Repeatability of each extraction method in terms of bacterial community assessed by multivariate homogeneity analysis of group dispersions for *A. cephalotes* (blue) and *P. penetrator* (yellow)

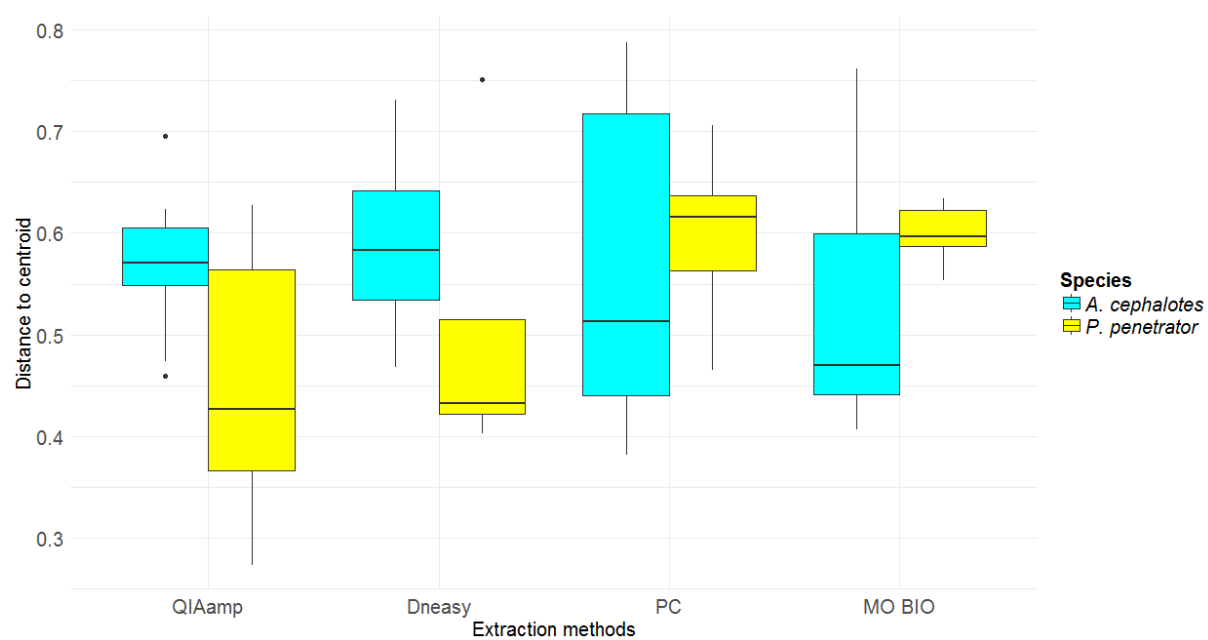


Figure S3. Bacterial clades log2 fold-change ratio between DNA extraction methods. This analysis allows to assess potential biases of extraction methods toward particular bacterial clades for *A. cephalotes* (a) and *P. penetrator* (b). Significance of change was assessed with a Mann-Whitney test with *fd*r p-value correction.

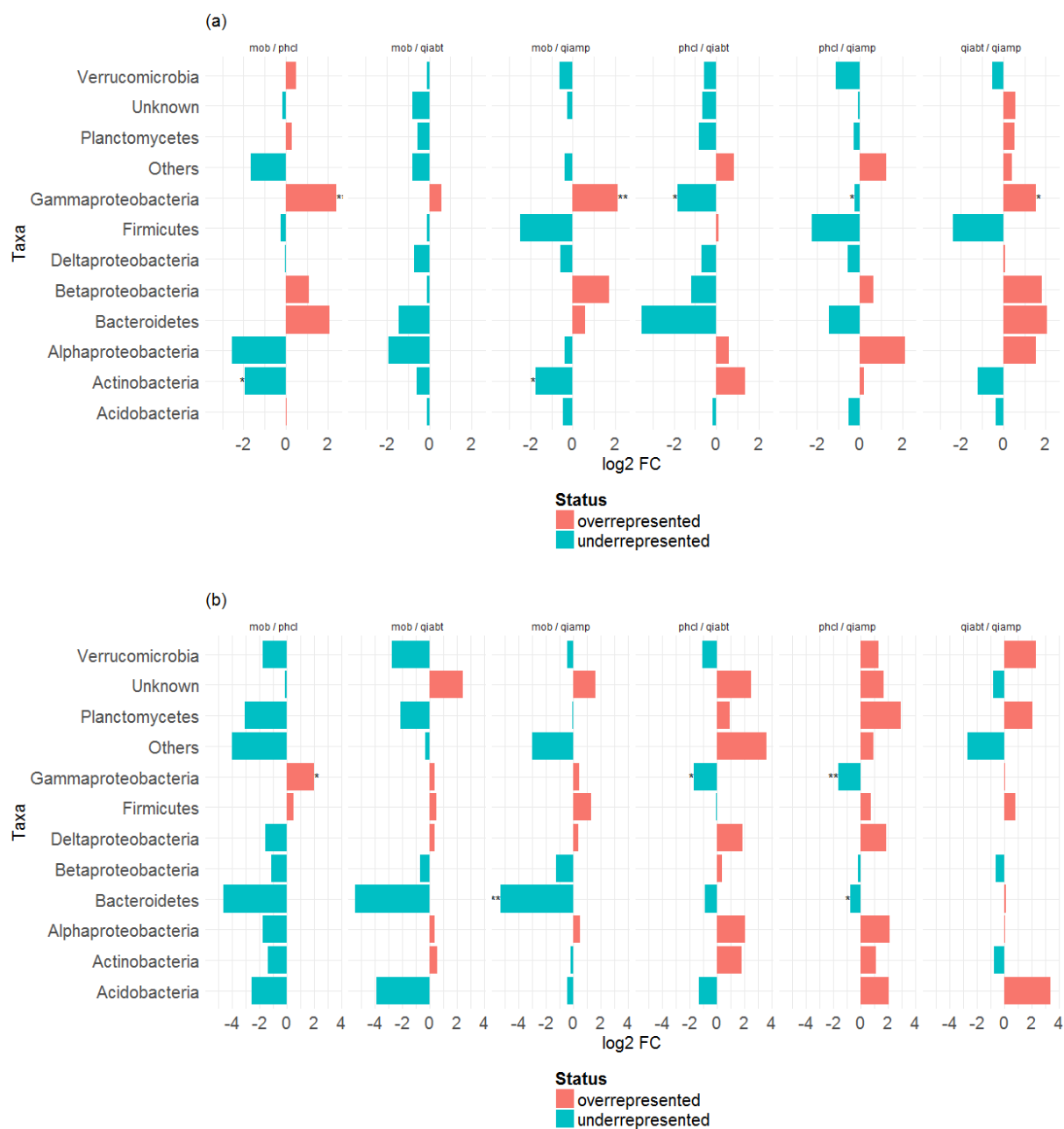


Figure S4. Effect of the number of individuals and individual body size on initial DNA concentration (a,d), number of reads (b,e) and OTUs (c,f). (a-c) Correspond to the comparisons of samples comprising one individual vs 10 individuals for both *A. cephalotes* and *P. penetrator*. (d-f) Correspond to comparison of samples made of pools of individuals of different body size in *A. cephalotes* (d-f).

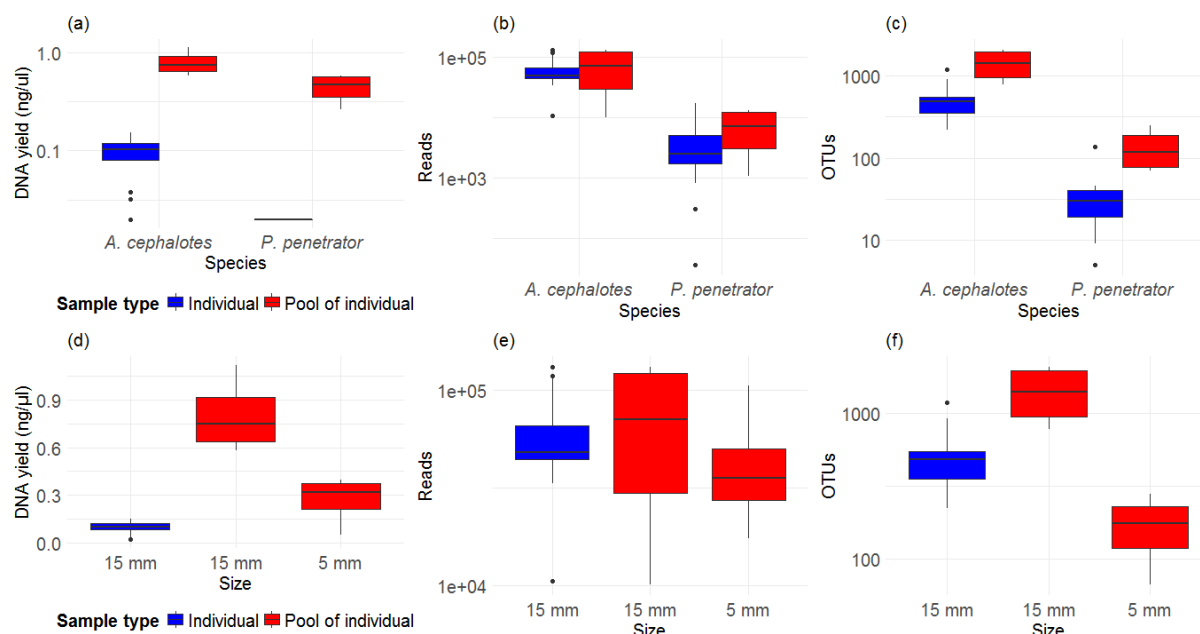


Figure S5. Rarefaction curves for averaged communities from single individual vs pools of 10 individuals for *A. cephalotes* and *P. penetrator* (a), and from pools of individuals differing in their body size for *A. cephalotes* (b).

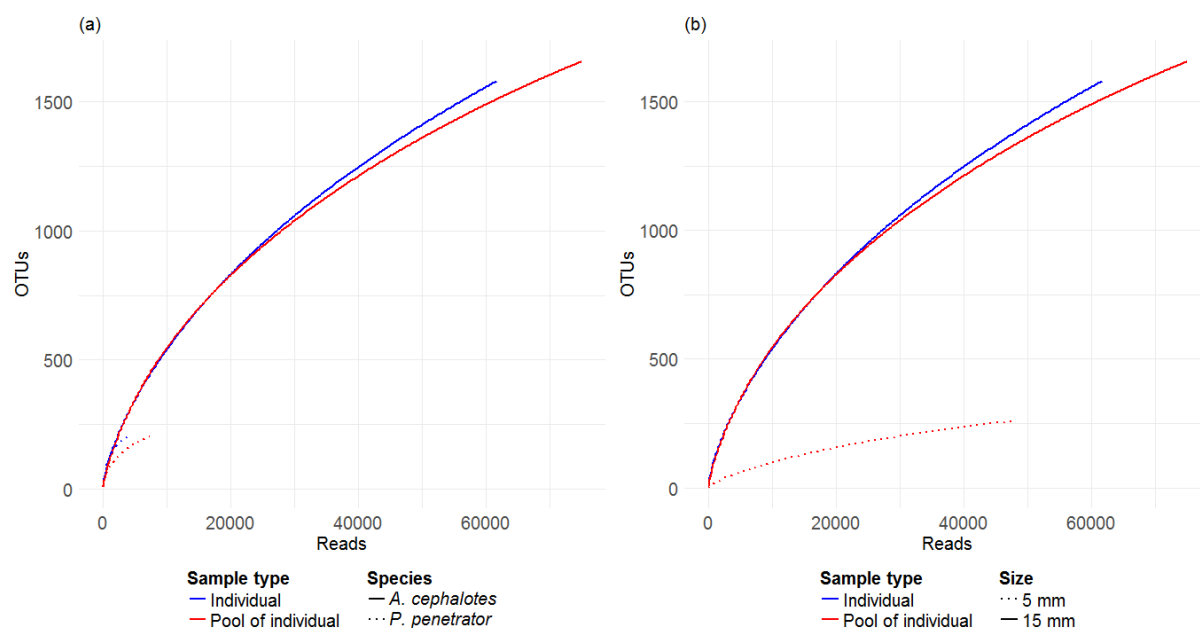


Figure S6. Repeatability of bacterial community composition when retrieved from one or pools of individuals or from individuals of different body size. The plot corresponds to an analysis of multivariate homogeneity of group dispersions.

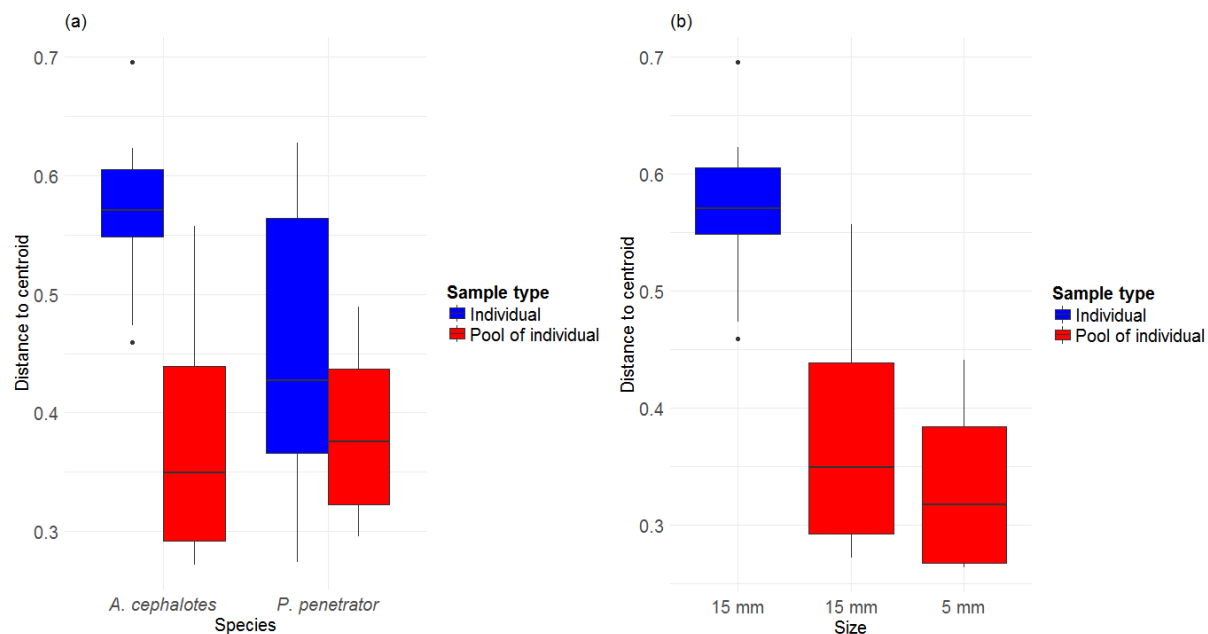


Table S1. Tukey post-hoc comparisons of the effects of DNA extraction methods and species on the proportion of (a) artifactual OTUs in samples, (b) contaminants OTUs in samples and (c) false positive in samples.

Linear hypotheses	(a)			(b)			(c)		
	Estimate	Pr(> t)		Estimate	Pr(> t)		Estimate	Pr(> t)	
PC <i>A.cephalotes</i> - MO BIO <i>A.cephalotes</i> = 0	0.08	0.77		0.28	0.03	*	0.05	0.90	
Dneasy <i>A.cephalotes</i> - MO BIO <i>A.cephalotes</i> = 0	-0.68	0.01	*	-0.04	0.74		1.79	3.44E-04	***
QIAamp <i>A.cephalotes</i> - MO BIO <i>A.cephalotes</i> = 0	-0.39	0.19		0.05	0.74		0.22	0.72	
MO BIO <i>P.penetrator</i> - MO BIO <i>A.cephalotes</i> = 0	-0.22	0.47		0.04	0.74		1.87	2.13E-04	***
PC <i>P.penetrator</i> - MO BIO <i>A.cephalotes</i> = 0	-0.07	0.77		0.76	8.33E-09	***	0.92	0.07	.
Dneasy <i>P.penetrator</i> - MO BIO <i>A.cephalotes</i> = 0	-1.03	5.73E-04	***	0.11	0.56		2.69	2.38E-06	***
QIAamp <i>P.penetrator</i> - MO BIO <i>A.cephalotes</i> = 0	-0.15	0.61		0.59	3.01E-06	***	0.38	0.49	
Dneasy <i>A.cephalotes</i> - PC <i>A.cephalotes</i> = 0	-0.76	0.01	**	-0.33	0.01	*	1.73	6.09E-04	***
QIAamp <i>A.cephalotes</i> - PC <i>A.cephalotes</i> = 0	-0.47	0.10	.	-0.23	0.09	.	0.17	0.77	
MO BIO <i>P.penetrator</i> - PC <i>A.cephalotes</i> = 0	-0.30	0.33		-0.24	0.08	.	1.82	3.44E-04	***
PC <i>P.penetrator</i> - PC <i>A.cephalotes</i> = 0	-0.15	0.61		0.47	2.64E-04	***	0.87	0.09	.
Dneasy <i>P.penetrator</i> - PC <i>A.cephalotes</i> = 0	-1.11	0.00	***	-0.18	0.26		2.63	3.25E-06	***
QIAamp <i>P.penetrator</i> - PC <i>A.cephalotes</i> = 0	-0.23	0.46		0.31	0.02	*	0.33	0.57	
QIAamp <i>A.cephalotes</i> - Dneasy <i>A.cephalotes</i> = 0	0.29	0.33		0.09	0.56		-1.56	1.34E-03	**
MO BIO <i>P.penetrator</i> - Dneasy <i>A.cephalotes</i> = 0	0.46	0.10	.	0.09	0.58		0.08	0.88	
PC <i>P.penetrator</i> - Dneasy <i>A.cephalotes</i> = 0	0.61	0.03	*	0.80	2.14E-09	***	-0.87	0.09	.
Dneasy <i>P.penetrator</i> - Dneasy <i>A.cephalotes</i> = 0	-0.35	0.29		0.15	0.35		0.90	0.10	.
QIAamp <i>P.penetrator</i> - Dneasy <i>A.cephalotes</i> = 0	0.52	0.06	.	0.64	7.21E-07	***	-1.40	3.92E-03	**
MO BIO <i>P.penetrator</i> - QIAamp <i>A.cephalotes</i> = 0	0.17	0.60		-0.01	0.95		1.65	8.31E-04	***
PC <i>P.penetrator</i> - QIAamp <i>A.cephalotes</i> = 0	0.32	0.29		0.70	4.44E-08	***	0.70	0.16	
Dneasy <i>P.penetrator</i> - QIAamp <i>A.cephalotes</i> = 0	-0.64	0.03	*	0.05	0.74		2.46	7.31E-06	***
QIAamp <i>P.penetrator</i> - QIAamp <i>A.cephalotes</i> = 0	0.24	0.46		0.54	1.73E-05	***	0.16	0.77	
PC <i>P.penetrator</i> - MO BIO <i>P.penetrator</i> = 0	0.15	0.61		0.71	4.34E-08	***	-0.95	0.07	.
Dneasy <i>P.penetrator</i> - MO BIO <i>P.penetrator</i> = 0	-0.81	0.01	**	0.06	0.74		0.82	0.14	
QIAamp <i>P.penetrator</i> - MO BIO <i>P.penetrator</i> = 0	0.06	0.77		0.55	1.50E-05	***	-1.48	2.27E-03	**
Dneasy <i>P.penetrator</i> - PC <i>P.penetrator</i> = 0	-0.96	1.13E-03	**	-0.65	2.88E-06	***	1.77	9.02E-04	***
QIAamp <i>P.penetrator</i> - PC <i>P.penetrator</i> = 0	-0.08	0.77		-0.16	0.26		-0.54	0.30	
QIAamp <i>P.penetrator</i> - Dneasy <i>P.penetrator</i> = 0	0.87	0.00	**	0.49	3.86E-04	***	-2.30	2.52E-05	***

Table S2. Table of contaminant OTUs identified in PCR controls (PCR), in tweezer controls (tweezer) and in DNA extraction controls

OTU identity	Class name	Lowest validated name	Sample type (number of reads)
DefAnts1_BA_seq_125273	Peptoniphilus	Clostridia	PCR (8)
DefAnts1_BA_seq_357561	Halobacteria	Halobacteria	PCR (4)
DefAnts1_BA_seq_258891	Bacteria	Opitutae	PCR (4)
DefAnts1_BA_seq_48921	Kocuria	Actinobacteria	PCR (33)
DefAnts1_BA_seq_217198	Geobacillus	Bacilli	PCR (55)
DefAnts1_BA_seq_0055	Bacteroidia	Bacteroides	tweezer (8210)
DefAnts1_BA_seq_0329	Sphingobacteriia	Pedobacter	tweezer (9368)
DefAnts1_BA_seq_0322	Gammaproteobacteria	Halomonas	tweezer (907)
DefAnts1_BA_seq_5606	Gammaproteobacteria	Halomonas	tweezer (1177)
DefAnts1_BA_seq_0401	Bacilli	Alkalibacterium	tweezer (26164)
DefAnts1_BA_seq_1796	Sphingobacteriia	Pedobacter	tweezer (893)
DefAnts1_BA_seq_0529	Gammaproteobacteria	Shewanella	tweezer (5812)
DefAnts1_BA_seq_137153	-	Bacteroidetes	Mb (16)
DefAnts1_BA_seq_34432	Bacilli	Bacilli	Mb (34)
DefAnts1_BA_seq_294686	Actinobacteria	Coriobacteriaceae	Mb (5)
DefAnts1_BA_seq_43419	Flavobacteriia	Flavobacterium	PC (4)
DefAnts1_BA_seq_95529	Clostridia	Anaerococcus	PC (26)
DefAnts1_BA_seq_17979	-	Proteobacteria	PC (7)
DefAnts1_BA_seq_21541	Gammaproteobacteria	Salinisphaera	DNeasy (5)
DefAnts1_BA_seq_29037	Bacilli	Lactobacillus	DNeasy (33), Mb(1)
DefAnts1_BA_seq_4397	Epsilonproteobacteria	Arcobacter	DNeasy (58), QIAamp (13)
DefAnts1_BA_seq_97666	Gammaproteobacteria	Pseudoxanthomonas	DNeasy (22)
DefAnts1_BA_seq_522813	-	Bacteria	DNeasy (2)
DefAnts1_BA_seq_57575	Actinobacteria	Solirubrobacter	DNeasy (16)
DefAnts1_BA_seq_59025	-	Bacteria	DNeasy (1)
DefAnts1_BA_seq_240789	Gammaproteobacteria	Xanthomonadaceae	DNeasy (3)
DefAnts1_BA_seq_253395	-	Bacteria	DNeasy (1)
DefAnts1_BA_seq_5864	Gammaproteobacteria	Vibrionaceae	DNeasy (159), QIAamp(114)
DefAnts1_BA_seq_477795	Actinobacteria	Actinomycetales	DNeasy (1), QIAamp(1)
DefAnts1_BA_seq_74765	Alphaproteobacteria	Pseudochrobactrum	DNeasy (6), QIAamp(19)
DefAnts1_BA_seq_3811	-	Bacteria	QIAamp (59)
DefAnts1_BA_seq_348517	-	Candidatus	QIAamp (2)
DefAnts1_BA_seq_9136	Alphaproteobacteria	Sphingomonas	QIAamp (79)
DefAnts1_BA_seq_2596	Actinobacteria	Corynebacterium	DNeasy (59), QIAamp (265)
DefAnts1_BA_seq_201154	Alphaproteobacteria	Rhizobiales	QIAamp (1)
DefAnts1_BA_seq_33643	Actinobacteria	Actinomyces	Mb (3), QIAamp (30)
DefAnts1_BA_seq_41767	Deinococci	Deinococcus	DNeasy (7), QIAamp (31)
DefAnts1_BA_seq_152198	Alphaproteobacteria	Amaricoccus	QIAamp (7)
DefAnts1_BA_seq_0620	Gammaproteobacteria	Psychrobacter	DNeasy (38), QIAamp (363)
DefAnts1_BA_seq_121214	Clostridia	Lutispora	QIAamp (11)
DefAnts1_BA_seq_41187	Cytophagia	Hymenobacter	QIAamp (10)
DefAnts1_BA_seq_250279	Actinobacteria	Acidimicrobiales	QIAamp (3)
DefAnts1_BA_seq_29700	Negativicutes	Veillonellaceae	QIAamp (57)
DefAnts1_BA_seq_2509	Bacilli	Facklamia	QIAamp (87)

DefAnts1_BA_seq_12796	Bacilli	Facklamia	QIAamp (33)
DefAnts1_BA_seq_1828	Bacilli	Bacillales	QIAamp (51)
DefAnts1_BA_seq_270057	Gammaproteobacteria	Gammaproteobacteria	QIAamp (17)
DefAnts1_BA_seq_154755	Bacteroidia	Dysgonomonas	QIAamp (3)
DefAnts1_BA_seq_33545	Betaproteobacteria	Betaproteobacteria	QIAamp (1)
DefAnts1_BA_seq_118064	Actinobacteria	Cellulomonas	QIAamp (23)
DefAnts1_BA_seq_15491	Erysipelotrichia	Erysipelothrix	QIAamp (49)
DefAnts1_BA_seq_97889	Betaproteobacteria	Oxalobacteraceae	QIAamp (22)
DefAnts1_BA_seq_4297	Gammaproteobacteria	Moraxellaceae	Mb (34), PC (49), QIAamp (118)
DefAnts1_BA_seq_84171	Alphaproteobacteria	Rhizobiales	QIAamp (13)
DefAnts1_BA_seq_43619	Actinobacteria	Solirubrobacterales	QIAamp (43)
DefAnts1_BA_seq_141902	Actinobacteria	Actinomycetales	QIAamp (25)
DefAnts1_BA_seq_4016	Bacilli	Lactococcus	QIAamp (59)
DefAnts1_BA_seq_48558	Actinobacteria	Brachybacterium	QIAamp (36)
DefAnts1_BA_seq_71505	Alphaproteobacteria	Sphingomonadaceae	QIAamp (24)
DefAnts1_BA_seq_46277	Flavobacteriia	Flavobacteriaceae	QIAamp (37)
DefAnts1_BA_seq_2773	Bacteroidia	Bacteroides	DNeasy (2), QIAamp (103)
DefAnts1_BA_seq_61066	-	Bacteroidetes	QIAamp (49)
DefAnts1_BA_seq_57769	Betaproteobacteria	Burkholderiales	QIAamp (4)
DefAnts1_BA_seq_3515	Gammaproteobacteria	Halomonas	DNeasy (16), QIAamp (268)
DefAnts1_BA_seq_36051	Clostridia	Tissierella	QIAamp (5)
DefAnts1_BA_seq_12600	Alphaproteobacteria	Acetobacteraceae	QIAamp (75)
DefAnts1_BA_seq_8437	Alphaproteobacteria	Rhizobium	QIAamp (24)
DefAnts1_BA_seq_17634	Betaproteobacteria	Acidovorax	QIAamp (29)
DefAnts1_BA_seq_37244	Actinobacteria	Propionibacteriaceae	QIAamp (35)
DefAnts1_BA_seq_28693	Cytophagia	Cytophagales	QIAamp (12)
DefAnts1_BA_seq_86561	Alphaproteobacteria	Sphingomonadaceae	QIAamp (28)
DefAnts1_BA_seq_119101	Flavobacteriia	Flavobacteriaceae	QIAamp (13)
DefAnts1_BA_seq_5166	Actinobacteria	Actinomycetales	QIAamp (229)
DefAnts1_BA_seq_71776	-	Proteobacteria	QIAamp (5)
DefAnts1_BA_seq_33683	Actinobacteria	Corynebacterium	QIAamp (28)
DefAnts1_BA_seq_28570	Alphaproteobacteria	Devosia	QIAamp (71)
DefAnts1_BA_seq_110405	Gammaproteobacteria	Rheinheimera	QIAamp (13)
DefAnts1_BA_seq_26154	Alphaproteobacteria	Acetobacteraceae	QIAamp (9)
DefAnts1_BA_seq_11939	Gammaproteobacteria	Legionella	QIAamp (21)
DefAnts1_BA_seq_70813	Cytophagia	Cytophagales	QIAamp (3)
DefAnts1_BA_seq_419585	Flavobacteriia	Chryseobacterium	QIAamp (2)
DefAnts1_BA_seq_132706	Actinobacteria	Demequina	QIAamp (12)
DefAnts1_BA_seq_72785	Alphaproteobacteria	Sphingomonas	QIAamp (14)
DefAnts1_BA_seq_181378	Alphaproteobacteria	Alphaproteobacteria	QIAamp (1)
DefAnts1_BA_seq_20581	Actinobacteria	Propionibacterium	QIAamp (80)
DefAnts1_BA_seq_124087	Gammaproteobacteria	Legionella	QIAamp (28)
DefAnts1_BA_seq_8245	Actinobacteria	Dermacoccus	QIAamp (42)
DefAnts1_BA_seq_20997	Actinobacteria	Rothia	DNeasy (8), QIAamp (57)
DefAnts1_BA_seq_57675	Gammaproteobacteria	Gammaproteobacteria	QIAamp (30)
DefAnts1_BA_seq_57337	Alphaproteobacteria	Sphingomonas	QIAamp (21)
DefAnts1_BA_seq_1474	Bacilli	Lactobacillus	PC (46), QIAamp (42)
DefAnts1_BA_seq_45198	Actinobacteria	Actinomycetales	QIAamp (35)

DefAnts1_BA_seq_34856	Gammaproteobacteria	Rheinheimera	QIAamp (3)
DefAnts1_BA_seq_33122	Gammaproteobacteria	Xanthomonadaceae	QIAamp (97)
DefAnts1_BA_seq_216441	-	Candidatus	QIAamp (21)
DefAnts1_BA_seq_1013	-	Candidatus	QIAamp (80)
DefAnts1_BA_seq_21746	Betaproteobacteria	Pusillimonas	QIAamp (17)
DefAnts1_BA_seq_44529	Gammaproteobacteria	Legionella	QIAamp (38)
DefAnts1_BA_seq_112970	-	Bacteria	QIAamp (15)
DefAnts1_BA_seq_24501	Alphaproteobacteria	Paracoccus	QIAamp (51)
DefAnts1_BA_seq_131482	Gammaproteobacteria	Halomonadaceae	QIAamp (22)
DefAnts1_BA_seq_5128	Actinobacteria	Actinomycetales	QIAamp (31)
DefAnts1_BA_seq_23330	Alphaproteobacteria	Methylobacterium	QIAamp (54)
DefAnts1_BA_seq_68058	Actinobacteria	Kocuria	QIAamp (55)
DefAnts1_BA_seq_63005	Alphaproteobacteria	Microvirga	PC (15)
DefAnts1_BA_seq_1600	Alphaproteobacteria	Roseomonas	PC (568), QIAamp (14)
DefAnts1_BA_seq_18361	Actinobacteria	Ilumatobacter	PC (32)
DefAnts1_BA_seq_1076	Alphaproteobacteria	Porphyrobacter	PC (302)
DefAnts1_BA_seq_0374	Alphaproteobacteria	Methylobacterium	PC (143)
DefAnts1_BA_seq_19961	Actinobacteria	Pseudonocardia	PC (131)
DefAnts1_BA_seq_55953	Cytophagia	Fibrisoma	PC (8)
DefAnts1_BA_seq_6548	Acidobacteria_Gp4	Acidobacteria_Gp4	PC (17)
DefAnts1_BA_seq_61034	-	Bacteria	PC (43)
DefAnts1_BA_seq_116407	Actinobacteria	Actinomycetales	PC (13)
DefAnts1_BA_seq_2279	Alphaproteobacteria	Xanthobacteraceae	PC (76)
DefAnts1_BA_seq_29676	Bacilli	Lactobacillales	PC (117)
DefAnts1_BA_seq_23576	Bacteroidia	Bacteroides	QIAamp (20)
DefAnts1_BA_seq_21404	Betaproteobacteria	Comamonadaceae	QIAamp (28)
DefAnts1_BA_seq_5036	Actinobacteria	Janibacter	QIAamp (15)
DefAnts1_BA_seq_8311	Alphaproteobacteria	Bosea	QIAamp (44)
DefAnts1_BA_seq_11427	Negativicutes	Zymophilus	QIAamp (37)
DefAnts1_BA_seq_32576	Betaproteobacteria	Oxalobacteraceae	QIAamp (51)
DefAnts1_BA_seq_1754	Alphaproteobacteria	Methylobacterium	QIAamp (44)
DefAnts1_BA_seq_16877	Alphaproteobacteria	Sphingomonas	QIAamp (14)
DefAnts1_BA_seq_71693	Actinobacteria	Actinobacteria	PC (22)
DefAnts1_BA_seq_255538	Planctomycetia	Gemmata	QIAamp (1)
DefAnts1_BA_seq_34940	Bacilli	Bacillaceae	QIAamp (15)
DefAnts1_BA_seq_143657	Gammaproteobacteria	Xanthomonadaceae	QIAamp (56)
DefAnts1_BA_seq_171915	-	Bacteroidetes	DNeasy (12)
DefAnts1_BA_seq_47494	Mollicutes	Acholeplasma	QIAamp (8)
DefAnts1_BA_seq_334331	Sphingobacteriia	Chitinophagaceae	QIAamp (4)
DefAnts1_BA_seq_318724	Alphaproteobacteria	Rhizobiales	QIAamp (1)
DefAnts1_BA_seq_73807	Gammaproteobacteria	Halomonas	QIAamp (15)
DefAnts1_BA_seq_11295	Flavobacteriia	Flavobacterium	QIAamp (56)
DefAnts1_BA_seq_7405	Alphaproteobacteria	Methylobacterium	QIAamp (34)
DefAnts1_BA_seq_204674	Flavobacteriia	Chryseobacterium	Mb (3)
DefAnts1_BA_seq_64849	Actinobacteria	Micrococcaceae	Mb (18)
DefAnts1_BA_seq_24662	Planctomycetia	Planctomycetaceae	QIAamp (1)
DefAnts1_BA_seq_32022	Bacilli	Paenibacillus	QIAamp (26)
DefAnts1_BA_seq_47468	Alphaproteobacteria	Alphaproteobacteria	QIAamp (7)

DefAnts1_BA_seq_73052	Bacilli	Bacilli	QIAamp (12)
DefAnts1_BA_seq_44008	Flavobacteriia	Flavobacterium	QIAamp (14)
DefAnts1_BA_seq_82789	Actinobacteria	Nocardiopsis	Mb (4)
DefAnts1_BA_seq_1765	Alphaproteobacteria	Devosia	QIAamp (61)
DefAnts1_BA_seq_37887	Actinobacteria	Nesterenkonia	QIAamp (10)
DefAnts1_BA_seq_119719	Chlamydiia	Parachlamydiaceae	QIAamp (7)
DefAnts1_BA_seq_245952	-	Bacteria	QIAamp (2)

Table S3. Tukey post-hoc comparisons of the effects of DNA extraction methods and species on the number of (a) reads in samples and (b) OTUs in samples.

Linear hypotheses	(a)			(b)		
	Estimate	Pr(> t)		Estimate	Pr(> t)	
Dneasy <i>A.cephalotes</i> - QIAamp <i>A.cephalotes</i> = 0	-1.53	0.02	*	-0.80	0.05	.
PC <i>A.cephalotes</i> - QIAamp <i>A.cephalotes</i> = 0	-0.04	0.95		-1.04	0.01	*
MO BIO <i>A.cephalotes</i> - QIAamp <i>A.cephalotes</i> = 0	-0.39	0.59		-0.41	0.34	
QIAamp <i>P.penetrator</i> - QIAamp <i>A.cephalotes</i> = 0	-2.74	2.21E-05	***	-2.70	3.62E-11	***
Dneasy <i>P.penetrator</i> - QIAamp <i>A.cephalotes</i> = 0	-5.76	<2E-16	***	-4.69	2.00E-16	***
PC <i>P.penetrator</i> - QIAamp <i>A.cephalotes</i> = 0	-3.82	2.75E-09	***	-3.64	2.00E-16	***
MO BIO <i>P.penetrator</i> - QIAamp <i>A.cephalotes</i> = 0	-2.63	4.19E-05	***	-2.86	2.82E-12	***
PC <i>A.cephalotes</i> - Dneasy <i>A.cephalotes</i> = 0	1.49	0.03	*	-0.24	0.58	
MO BIO <i>A.cephalotes</i> - Dneasy <i>A.cephalotes</i> = 0	1.14	0.08	.	0.40	0.34	
QIAamp <i>P.penetrator</i> - Dneasy <i>A.cephalotes</i> = 0	-1.21	0.07	.	-1.89	3.61E-06	***
Dneasy <i>P.penetrator</i> - Dneasy <i>A.cephalotes</i> = 0	-4.23	3.81E-11	***	-3.89	2.00E-16	***
PC <i>P.penetrator</i> - Dneasy <i>A.cephalotes</i> = 0	-2.29	3.58E-04	***	-2.83	6.80E-12	***
MO BIO <i>P.penetrator</i> - Dneasy <i>A.cephalotes</i> = 0	-1.10	0.09	.	-2.05	5.42E-07	***
MO BIO <i>A.cephalotes</i> - PC <i>A.cephalotes</i> = 0	-0.35	0.62		0.64	0.13	
QIAamp <i>P.penetrator</i> - PC <i>A.cephalotes</i> = 0	-2.70	4.09E-05	***	-1.66	7.11E-05	***
Dneasy <i>P.penetrator</i> - PC <i>A.cephalotes</i> = 0	-5.72	<2E-16	***	-3.65	2.00E-16	***
PC <i>P.penetrator</i> - PC <i>A.cephalotes</i> = 0	-3.78	7.55E-09	***	-2.60	7.40E-10	***
MO BIO <i>P.penetrator</i> - PC <i>A.cephalotes</i> = 0	-2.59	7.54E-05	***	-1.82	1.40E-05	***
QIAamp <i>P.penetrator</i> - MO BIO <i>A.cephalotes</i> = 0	-2.36	2.49E-04	***	-2.29	1.93E-08	***
Dneasy <i>P.penetrator</i> - MO BIO <i>A.cephalotes</i> = 0	-5.38	<2E-16	***	-4.29	2.00E-16	***
PC <i>P.penetrator</i> - MO BIO <i>A.cephalotes</i> = 0	-3.43	9.00E-08	***	-3.23	4.14E-15	***
MO BIO <i>P.penetrator</i> - MO BIO <i>A.cephalotes</i> = 0	-2.25	4.44E-04	***	-2.45	1.88E-09	***
Dneasy <i>P.penetrator</i> - QIAamp <i>P.penetrator</i> = 0	-3.02	2.70E-06	***	-2.00	3.61E-06	***
PC <i>P.penetrator</i> - QIAamp <i>P.penetrator</i> = 0	-1.08	0.09	.	-0.94	0.03	*
MO BIO <i>P.penetrator</i> - QIAamp <i>P.penetrator</i> = 0	0.11	0.89		-0.16	0.69	
PC <i>P.penetrator</i> - Dneasy <i>P.penetrator</i> = 0	1.94	2.55E-03	**	1.06	0.02	*
MO BIO <i>P.penetrator</i> - Dneasy <i>P.penetrator</i> = 0	3.13	1.20E-06	***	1.84	2.02E-05	***
MO BIO <i>P.penetrator</i> - PC <i>P.penetrator</i> = 0	1.19	0.07	.	0.78	0.07	.

Annexe 7 : Table des OTUs communs aux deux espèces de fourmis (remplissage orange) et aux deux espèces de fourmis et au nid (remplissage jaune) pour les trois sites

Code de l'OTU	Nom classe	Nom scientifique juste	Kaw	Petit Saut	Nouragues
seq_13561		candidate division WPS-1	C. fem&C. lev&Nid	C. lev	C. fem
seq_3529		Candidatus Saccharibacteria	C. fem&C. lev&Nid	C. lev&Nid	C. fem&C. lev&Nid
seq_2314		Bacteria	C. fem&C. lev&Nid	C. lev	Rare
seq_76694		Armatimonadetes	C. fem&C. lev&Nid	C. lev	Rare
seq_3646		Acidobacteria_Gp1	C. fem&C. lev&Nid	C. lev&Nid	Rare
seq_2435		Acidobacteria_Gp1	C. fem&C. lev	Rare	-
seq_0360	Acidobacteria_Gp1	Acidobacteria_Gp1	C. fem&C. lev&Nid	C. fem&C. lev&Nid	C. fem&C. lev&Nid
seq_12582		Acidobacteria_Gp1	C. lev	C. fem&C. lev	C. fem
seq_0451	Acidobacteria_Gp2	Acidobacteria_Gp2	C. fem&C. lev	C. lev	-
seq_7509		Acidobacteria_Gp3	C. fem&C. lev&Nid	C. lev	Rare
seq_2727	Acidobacteria_Gp3	Acidobacteria_Gp3	C. fem&C. lev	C. lev	Rare
seq_766391		Nocardiaceae	C. fem&C. lev&Nid	C. lev&Nid	C. fem
seq_31154		Pseudonocardia	C. fem	C. fem&C. lev&Nid	C. fem
seq_23692		Actinobacteria	Nid	Nid	C. fem&C. lev
seq_10428		Microbacterium	C. fem&C. lev&Nid	C. fem&C. lev&Nid	C. fem&C. lev
seq_28448		Acidimicrobiales	C. fem&C. lev&Nid	C. lev&Nid	C. fem
seq_24313		Actinomycetales	Rare	C. fem&C. lev	Rare
seq_3381		Nakamurella	C. fem&C. lev&Nid	C. fem&C. lev&Nid	C. fem&C. lev
seq_0944		Mycobacterium	C. fem&C. lev&Nid	C. fem&C. lev&Nid	C. fem&C. lev&Nid
seq_16030		Streptacidiphilus	C. fem&C. lev&Nid	Rare	Rare
seq_9627		Nocardiaceae	C. fem&C. lev	Rare	Rare
seq_15517		Jatrophihabitans	C. fem&C. lev&Nid	C. lev&Nid	C. fem&C. lev
seq_0309	Actinobacteria	Propionibacterium	Rare	C. fem&C. lev	C. lev
seq_0905		Jatrophihabitans	C. fem&C. lev&Nid	C. fem&C. lev&Nid	C. fem&C. lev&Nid
seq_0004		Pseudonocardia	C. fem&C. lev	Rare	Rare
seq_1630		Solirubrobacterales	C. fem&C. lev&Nid	C. lev&Nid	Rare
seq_3052		Jatrophihabitans	Rare	C. fem&C. lev	C. fem&C. lev
seq_0191		Nocardiaceae	C. fem&C. lev&Nid	C. lev	C. fem&C. lev
seq_1124		Actinomycetales	C. fem&C. lev	Rare	-
seq_1463		Nocardiaceae	C. fem&C. lev&Nid	Rare	C. fem
seq_0405		Actinomycetales	C. fem&C. lev	C. lev&Nid	Rare
seq_6033		Actinomycetales	Rare	Rare	C. fem&C. lev
seq_7380		Microbacterium	Nid	Rare	C. fem&C. lev&Nid
seq_0218		Bradyrhizobiaceae	C. fem&C. lev	C. lev&Nid	C. lev
seq_10178		Bradyrhizobiaceae	C. lev&Nid	Rare	C. fem&C. lev
seq_22113		Rhizobiales	Rare	Rare	C. fem&C. lev
seq_0291		Rhizobiales	C. fem&C. lev&Nid	C. lev	C. fem&C. lev
seq_4210		Alphaproteobacteria_incertae_sedis	Rare	C. lev	C. fem&C. lev&Nid
seq_12905		Rhizobiales	C. fem&C. lev&Nid	C. lev	Rare
seq_3788		Acetobacteraceae	C. fem&C. lev	-	Rare
seq_3611		Sphingomonas	Rare	Rare	C. fem&C. lev
seq_6175	Alphaproteobacteria	Acetobacteraceae	C. fem&C. lev	C. lev&Nid	C. fem
seq_1649		Bradyrhizobium	C. fem&C. lev&Nid	C. fem&C. lev&Nid	C. fem&C. lev&Nid
seq_6099		Varilbacter	C. fem&C. lev&Nid	C. fem&C. lev&Nid	C. fem&C. lev
seq_0243		Phenylobacterium	C. fem&C. lev&Nid	C. lev	C. lev
seq_19085		Mesorhizobium	Rare	C. lev	C. fem&C. lev
seq_4817		Rhizobiales	C. fem&C. lev&Nid	C. lev	C. fem&C. lev
seq_26573		Alphaproteobacteria	Rare	Rare	C. fem&C. lev&Nid
seq_3605		Sphingomonas	Rare	C. fem&C. lev	Rare
seq_7327		Devosia	Rare	C. lev&Nid	C. fem&C. lev
seq_0440		Burkholderiales_incertae_sedis	C. fem&C. lev&Nid	C. fem&C. lev	C. fem&C. lev
seq_0197	Betaproteobacteria	Undibacterium	C. fem&C. lev	Rare	Rare
seq_0925		Betaproteobacteria	C. fem&C. lev&Nid	C. lev	C. fem&C. lev
seq_3402		Oxalobacteraceae	C. fem&C. lev&Nid	Rare	C. lev
seq_5029		Deltaproteobacteria	C. fem&C. lev&Nid	Rare	C. fem&C. lev
seq_14012		Labilithrix	C. fem&C. lev	C. lev	C. fem&C. lev
seq_37735	Deltaproteobacteria	Labilithrix	C. fem&C. lev&Nid	C. fem&C. lev	Rare
seq_7674		Myxococcales	C. fem&C. lev&Nid	C. lev	Rare
seq_0877		Gammaproteobacteria	C. fem&C. lev&Nid	C. lev	C. lev
seq_0421		Rhizobacter	C. fem&C. lev&Nid	C. lev	C. fem&C. lev
seq_31373		Pseudomonadales	C. fem&C. lev&Nid	Rare	C. fem
seq_2788		Gammaproteobacteria	C. fem&C. lev&Nid	C. lev	C. fem&C. lev
seq_24900	Planctomycetia	Planctomycetaceae	C. fem&C. lev	Rare	Rare
seq_31828		Chitinophagaceae	C. fem&C. lev&Nid	Rare	Rare
seq_821574	Sphingobacteriia	Chitinophagaceae	C. fem&C. lev	Rare	Rare
seq_4377		Chitinophagaceae	C. fem&C. lev	C. fem&C. lev	Rare
seq_9331	Subdivision3	Subdivision3	C. fem&C. lev	C. lev	C. fem

Bibliographie :

- Abdelmohsen UR, Bayer K, Hentschel U (2014) Diversity, abundance and natural products of marine sponge-associated actinomycetes. *Natural Product Reports*, **31**, 381–399.
- Abdelmohsen UR, Grkovic T, Balasubramanian S *et al.* (2015) Elicitation of secondary metabolism in actinomycetes. *Biotechnology Advances*, **33**, 798–811.
- Adermann K, John H, Ständker L, Forssmann W-G (2004) Exploiting natural peptide diversity: novel research tools and drug leads. *Current opinion in biotechnology*, **15**, 599–606.
- Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews*, **59**, 143–169.
- Ameur RM-B, Mellouli L, Chabchoub F, Fotso S, Bejar S (2004) Purification and structure elucidation of two biologically active molecules from a new isolated *Streptomyces* sp. US 24 strain. *Chemistry of Natural Compounds*, **40**, 510–513.
- Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, **26**, 32–46.
- Anderson MJ (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, **62**, 245–253.
- Baltz RH (2008) Renaissance in antibacterial discovery from actinomycetes. *Current Opinion in Pharmacology*, **8**, 557–563.
- Barka EA, Vatsa P, Sanchez L *et al.* (2016) Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews*, **80**, 1–43.
- de Bary A (1879) Die Erscheinung der Symbiose. *Karl J. Trubner*.
- Bass-Becking L (1934) Geobiologie of Inleiding tot de Milieukunde. *Van Stockum & Zoon*.
- Basset Y, Cizek L, Cuénoud P *et al.* (2012) Arthropod diversity in a tropical forest. *Science*, **338**, 1481–1484.

- Beale DJ, Karpe AV, Ahmed W (2016) Beyond Metabolomics: A Review of Multi-Omics-Based Approaches. In: *Microbial Metabolomics*, pp. 289–312. Springer.
- Belin-Depoux M (1991) Ecologie et evolution des jardins de fourmis en Guyane française. *Revue d'Ecologie*, 1–38.
- Bello JE, McElfresh JS, Millar JG (2015) Isolation and determination of absolute configurations of insect-produced methyl-branched hydrocarbons. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 1077–1082.
- Bérdy J (2005) Bioactive Microbial Metabolites. *The Journal of Antibiotics*, **58**, 1–26.
- Bertrand S, Bohni N, Schnee S *et al.* (2014) Metabolite induction via microorganism co-culture: a potential way to enhance chemical diversity for drug discovery. *Biotechnology Advances*, **32**, 1180–1204.
- Blair JMA, Webber MA, Baylay AJ, Ogbolu DO, Piddock LJV (2015) Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, **13**, 42–51.
- Blomquist GJ, Bagnères A-G (2010) *Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology*. Cambridge University Press.
- Borthwick AD (2012) 2,5-Diketopiperazines: Synthesis, Reactions, Medicinal Chemistry, and Bioactive Natural Products. *Chemical Reviews*, **112**, 3641–3716.
- Bot ANM, Ortius-Lechner D, Finster K, Maile R, Boomsma JJ (2002) Variable sensitivity of fungi and bacteria to compounds produced by the metapleural glands of leaf-cutting ants. *Insectes Sociaux*, **49**, 363–370.
- Bouslimani A, Porto C, Rath CM *et al.* (2015) Molecular cartography of the human skin surface in 3D. *Proceedings of the National Academy of Sciences*, **112**, 2120–2129.
- Boutin S, Sauvage C, Bernatchez L, Audet C, Derome N (2014) Inter individual variations of the fish skin microbiota: Host genetics basis of mutualism? *PLoS ONE*, **9**, 1–17.
- Boyer F, Mercier C, Bonin A *et al.* (2016) obitools: a unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, **16**, 176–182.

- Brownlie JC, Johnson KN (2009) Symbiont-mediated protection in insect hosts. *Trends in Microbiology*, **17**, 348–354.
- Burgess JG, Jordan EM, Bregu M, Mearns-Spragg A, Boyd KG (1999) Microbial antagonism: a neglected avenue of natural products research. *Journal of Biotechnology*, **70**, 27–32.
- Burgot G, Burgot J-L (2011) *Méthodes instrumentales d'analyse chimique et applications : Méthodes chromatographiques, électrophorèses, méthodes spectrales et méthodes thermiques*. Lavoisier.
- Carr G, Poulsen M, Klassen JL *et al.* (2012) Microtermolides A and B from Termite-Associated *Streptomyces* sp. and Structural Revision of Vinylamycin. *Organic Letters*, **14**, 2822–2825.
- Caruso T, Chan Y, Lacap DC *et al.* (2011) Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. *The ISME journal*, **5**, 1406–1413.
- Challinor VL, Bode HB (2015) Bioactive natural products from novel microbial sources. *Annals of the New York Academy of Sciences*, **1354**, 82–97.
- Chapela IH, Rehner SA, Schultz TR, Mueller UG (1994) Evolutionary history of the symbiosis between fungus-growing ants and their fungi. *Science (New York, N.Y.)*, **266**, 1691–1694.
- Charlop-Powers Z, Pregitzer CC, Lemetre C *et al.* (2016) Urban park soil microbiomes are a rich reservoir of natural product biosynthetic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 14811–14816.
- Chater KF, Biró S, Lee KJ, Palmer T, Schrempf H (2010) The complex extracellular biology of *Streptomyces*. *FEMS Microbiology Reviews*, **34**, 171–198.
- Chu J, Vila-Farres X, Inoyama D *et al.* (2016) Discovery of MRSA active antibiotics using primary sequence from the human microbiome. *Nature Chemical Biology*.

- Chun J, Lee J-H, Jung Y *et al.* (2007) EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences. *International Journal of Systematic and Evolutionary Microbiology*, **57**, 2259–2261.
- Clardy J, Fischbach M, Currie C (2009) The natural history of antibiotics. *Current biology : CB*, **19**, R437–R441.
- Claverías FP, Undabarrena A, González M, Seeger M, Cámara B (2015) Culturable diversity and antimicrobial activity of Actinobacteria from marine sediments in Valparaíso bay, Chile. *Frontiers in Microbiology*, **6**, 737.
- Clegg GA, Dole M (1971) Molecular beams of macroions. III. Zein and polyvinylpyrrolidone. *Biopolymers*, **10**, 821–826.
- Cogen AL, Nizet V, Gallo RL (2008) Skin microbiota: A source of disease or defence? *British Journal of Dermatology*, **158**, 442–455.
- Cohan FM (2002) What are Bacterial Species? *Annual Review of Microbiology*, **56**, 457–487.
- Corbara B, Dejean A, Orivel J (1999) Les «jardins de fourmis, une association plantes-fourmis originale. *L'Année Biologique*, **38**, 73–89.
- Currie CR, Mueller UG, Malloch D (1999a) The agricultural pathology of ant fungus gardens. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 7998–8002.
- Currie CR, Poulsen M, Mendenhall J, Boomsma JJ, Billen J (2006) Coevolved crypts and exocrine glands support mutualistic bacteria in fungus-growing ants. *Science (New York, N.Y.)*, **311**, 81–83.
- Currie CR, Scott JA, Summerbell RC, Malloch D (1999b) Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. *Nature*, **398**, 701–704.
- Curtis TP, Sloan WT (2004) Prokaryotic diversity and its limits: microbial community structure in nature and implications for microbial ecology. *Current Opinion in Microbiology*, **7**, 221–226.

- Dângelo RAC, de Souza DJ, Mendes TD, Couceiro J da C, Lucia TMCD (2016) Actinomycetes inhibit filamentous fungi from the cuticle of *Acromyrmex* leafcutter ants. *Journal of Basic Microbiology*, **56**, 229–237.
- Davies J, Davies D (2010) Origins and Evolution of Antibiotic Resistance. *Microbiology and Molecular Biology Reviews*, **74**, 417–433.
- Davis TS, Crippen TL, Hofstetter RW, Tomberlin JK (2013) Microbial Volatile Emissions as Insect Semiochemicals. *Journal of Chemical Ecology*, **39**, 840–859.
- Dejean A, Solano PJ, Ayroles J, Corbara B, Orivel J (2005) Arboreal ants build traps to capture prey. *Nature*, **434**, 973.
- Di Guida R, Engel J, Allwood JW *et al.* (2016) Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics: Official Journal of the Metabolomic Society*, **12**, 93.
- Dieter E (2013) The Epidemiology and Evolution of Symbionts with Mixed-Mode Transmission. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 623–643.
- Dieterle F, Ross A, Schlotterbeck G, Senn H (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Analytical Chemistry*, **78**, 4281–4290.
- Doroghazi JR, Metcalf WW (2013) Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes. *BMC Genomics*, **14**, 611.
- Dosmann A, Bahet N, Gordon DM (2016) Experimental modulation of external microbiome affects nestmate recognition in harvester ants (*Pogonomyrmex barbatus*). *PeerJ*, **4**, e1566.
- Douglas AE (2015) Multiorganismal insects: diversity and function of resident microorganisms. *Annual Review of Entomology*, **60**, 17–34.

- Douglas AE, Werren JH (2016) Holes in the Hologenome: Why Host-Microbe Symbioses Are Not Holobionts. *mBio*, **7**, e02099-15.
- Emery VJ, Tsutsui ND (2013) Recognition in a Social Symbiosis: Chemical Phenotypes and Nestmate Recognition Behaviors of Neotropical Parabiotic Ants. *PLoS ONE*, **8**.
- Emery VJ, Tsutsui ND (2016) Differential Sharing of Chemical Cues by Social Parasites Versus Social Mutualists in a Three-Species Symbiosis. *Journal of Chemical Ecology*, **42**, 277–285.
- Esling P, Lejzerowicz F, Pawlowski J (2015) Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Research*, **43**, 2513–2524.
- Fernández-Marín H, Zimmerman JK, Rehner SA, Wcislo WT (2006) Active use of the metapleural glands by ants in controlling fungal infection. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 1689–1695.
- Fierer N, Hamady M, Lauber CL, Knight R (2008) The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proceedings of the National Academy of Sciences*, 17994–17999.
- Fierer N, Strickland MS, Liptzin D, Bradford MA, Cleveland CC (2009) Global patterns in belowground communities. *Ecology Letters*, **12**, 1238–1249.
- Finlay BJ, Clarke KJ (1999) Ubiquitous dispersal of microbial species. *Nature*, **400**, 828–828.
- Fliegerova K, Tapio I, Bonin A *et al.* (2014) Effect of DNA extraction and sample preservation method on rumen bacterial population. *Anaerobe*, **29**, 80–84.
- Forel A (1898) La parabiose chez les fourmis. *Bulletin de la Société Vaudoise de Sciences Naturelles*, 380–384.
- Frank JA, Reich CI, Sharma S *et al.* (2008) Critical Evaluation of Two Primers Commonly Used for Amplification of Bacterial 16S rRNA Genes. *Applied and Environmental Microbiology*, **74**, 2461–2470.

- Franzosa EA, Hsu T, Sirota-Madi A *et al.* (2015) Sequencing and beyond: integrating molecular “omics” for microbial community profiling. *Nature Reviews Microbiology*, **13**, 360–372.
- Fuhrman JA (2009) Microbial community structure and its functional implications. *Nature*, **459**, 193–199.
- Gao H, Grischow S, Barke J *et al.* (2014) Filipins: the first antifungal “weed killers” identified from bacteria isolated from the trap-ant. *RSC Advances*, **4**, 57267–57270.
- Gefrides LA, Powell MC, Donley MA, Kahn R (2010) UV irradiation and autoclave treatment for elimination of contaminating DNA from laboratory consumables. *Forensic Science International: Genetics*, **4**, 89–94.
- Gerber NN (1967) Phenazines, phenoxazinones, and dioxopiperazines from *Streptomyces thioluteus*. *The Journal of Organic Chemistry*, **32**, 4055–4057.
- Ghoul M, Mitri S (2016) The Ecology and Evolution of Microbial Competition. *Trends in Microbiology*, **24**, 833–845.
- Goering AW, McClure RA, Doroghazi JR *et al.* (2016) Metabologenomics: Correlation of Microbial Gene Clusters with Metabolites Drives Discovery of a Nonribosomal Peptide with an Unusual Amino Acid Monomer. *ACS Central Science*, **2**, 99–108.
- González I, Lê Cao K-A, Melissa D, Déjean S (2013) Insightful graphical outputs to explore relationships between two “omics” data sets. *BioData Mining*, 5–19.
- Goodfellow M, Williams ST (1983) Ecology of actinomycetes. *Annual Review of Microbiology*, **37**, 189–216.
- Goodrich JK, Di Rienzi SC, Poole AC *et al.* (2014) Conducting a microbiome study. *Cell*, **158**, 250–262.
- Grice EA, Segre JA (2011) The skin microbiome. *Nature Reviews. Microbiology*, **9**, 244–253.
- Gromski PS, Xu Y, Kotze HL *et al.* (2014) Influence of missing values substitutes on multivariate analysis of metabolomics data. *Metabolites*, **4**, 433–452.

- Haeder S, Wirth R, Herz H, Spiteller D (2009) Candidicin-producing *Streptomyces* support leaf-cutting ants to protect their fungus garden against the pathogenic fungus *Escovopsis*. *Proceedings of the National Academy of Sciences*, **106**, 4742–4746.
- Hanbo Chen (2016) VennDiagram: Generate High-Resolution Venn and Euler Plots.
- Hanshew AS, McDonald BR, Díaz Díaz C *et al.* (2015) Characterization of actinobacteria associated with three ant-plant mutualisms. *Microbial Ecology*, **69**, 192–203.
- Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews. Microbiology*, **10**, 497–506.
- Harvey AL, Edrada-Ebel R, Quinn RJ (2015) The re-emergence of natural products for drug discovery in the genomics era. *Nature Reviews. Drug Discovery*, **14**, 111–129.
- Henrot M, Jean A, Peixoto PA, Maddaluno J, De Paolis M (2016) Flexible Total Synthesis of (±)-Aureothin, a Potent Antiproliferative Agent. *The Journal of Organic Chemistry*, **81**, 5190–5201.
- Hirata Y, Okuhara K, Naitô T (1954) A Toxic Nitro Compound from *Streptomyces thioluteus*. *Nature*, **173**, 1101–1101.
- Hölldobler B, Wilson EO (1990) *The Ants*. Harvard University Press.
- Hrydziuszko O, Viant MR (2012) Missing values in mass spectrometry based metabolomics: an undervalued step in the data processing pipeline. *Metabolomics*, **8**, 161–174.
- Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG, Qian K (2001) Kendrick Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra. *Analytical Chemistry*, **73**, 4676–4681.
- Hulcr J, Rountree NR, Diamond SE *et al.* (2012) Mycangia of ambrosia beetles host communities of bacteria. *Microbial Ecology*, **64**, 784–793.
- Hussa EA, Goodrich-Blair H (2013) It Takes a Village: Ecological and Fitness Impacts of Multipartite Mutualism. *Annual Review of Microbiology*, **67**, 161–178.

- Ishak HD, Plowes R, Sen R *et al.* (2011) Bacterial diversity in *Solenopsis invicta* and *Solenopsis geminata* ant colonies characterized by 16S amplicon 454 pyrosequencing. *Microbial Ecology*, **61**, 821–831.
- Jacques IB, Moutiez M, Witwinowski J *et al.* (2015) Analysis of 51 cyclodipeptide synthases reveals the basis for substrate specificity. *Nature Chemical Biology*, **11**, 721–727.
- Jami M, Ghanbari M, Kneifel W, Domig KJ (2015) Phylogenetic diversity and biological activity of culturable Actinobacteria isolated from freshwater fish gut microbiota. *Microbiological Research*, **175**, 6–15.
- Ju K-S, Gao J, Doroghazi JR *et al.* (2015) Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proceedings of the National Academy of Sciences*, **112**, 12175–12180.
- Juárez P, Chase J, Blomquist GJ (1992) A microsomal fatty acid synthetase from the integument of *Blattella germanica* synthesizes methyl-branched fatty acids, precursors to hydrocarbon and contact sex pheromone. *Archives of Biochemistry and Biophysics*, **293**, 333–341.
- Kaltenpoth M (2009) Actinobacteria as mutualists: general healthcare for insects? *Trends in Microbiology*, **17**, 529–535.
- Kaltenpoth M, Engl T (2014) Defensive microbial symbionts in Hymenoptera. *Functional Ecology*, **28**, 315–327.
- Kaltenpoth M, Yildirim E, Gürbüz MF, Herzner G, Strohm E (2012) Refining the roots of the beewolf-*Streptomyces* symbiosis: antennal symbionts in the rare genus *Philanthinus* (Hymenoptera, Crabronidae). *Applied and Environmental Microbiology*, **78**, 822–827.
- Kembel SW, O'Connor TK, Arnold HK *et al.* (2014) Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proceedings of the National Academy of Sciences*, **111**, 13715–13720.

- Kendrick E (1963) A Mass Scale Based on $\text{CH}_2 = 14.0000$ for High Resolution Mass Spectrometry of Organic Compounds. *Analytical Chemistry*, **35**, 2146–2154.
- Kessner D, Chambers M, Burke R, Agus D, Mallick P (2008) ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics (Oxford, England)*, **24**, 2534–2536.
- Kim O-S, Cho Y-J, Lee K *et al.* (2012) Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *International Journal of Systematic and Evolutionary Microbiology*, **62**, 716–721.
- Kingston W (2004) Streptomycin, Schatz v. Waksman, and the Balance of Credit for Discovery. *Journal of the History of Medicine and Allied Sciences*, **59**, 441–462.
- Knapp M, Hofreiter M (2010) Next generation sequencing of ancient DNA: Requirements, strategies and perspectives. *Genes*, **1**, 227–243.
- Knight R, Jansson J, Field D *et al.* (2012) Unlocking the potential of metagenomics through replicated experimental design. *Nature Biotechnology*, **30**, 513–520.
- Kroiss J, Kaltenpoth M, Schneider B *et al.* (2010) Symbiotic streptomycetes provide antibiotic combination prophylaxis for wasp offspring. *Nature Chemical Biology*, **6**, 261–263.
- Kueneman JG, Parfrey LW, Woodhams DC *et al.* (2014) The amphibian skin-associated microbiome across species, space and life history stages. *Molecular Ecology*, **23**, 1238–1250.
- Larsen A, Tao Z, Bullard SA, Arias CR (2013) Diversity of the skin microbiota of fishes: Evidence for host species specificity. *FEMS Microbiology Ecology*, **85**, 483–494.
- Lê Cao K-A, Boitard S, Besse P (2011) Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics*, **12**, 253.
- Lê Cao K-A, González I, Déjean S (2009) integrOmics: an R package to unravel relationships between two omics datasets. *Bioinformatics (Oxford, England)*, **25**, 2855–2856.

- Legendre P, Legendre LFJ (2012) *Numerical Ecology*. Elsevier.
- Lemaitre B, Nicolas E, Michaut L, Reichhart J-M, Hoffmann JA (1996) The Dorsoventral Regulatory Gene Cassette *spätzle/Toll/cactus* Controls the Potent Antifungal Response in *Drosophila* Adults. *Cell*, **86**, 973–983.
- Ley RE, Hamady M, Lozupone C *et al.* (2008) Evolution of mammals and their gut microbes. *Science (New York, N.Y.)*, **320**, 1647–1651.
- Li N, Korboukh VK, Krebs C, Bollinger JM (2010) Four-electron oxidation of p-hydroxylaminobenzoate to p-nitrobenzoate by a peroxodiferric complex in AurF from *Streptomyces thioluteus*. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 15722–15727.
- Ling LL, Schneider T, Peoples AJ *et al.* (2015) A new antibiotic kills pathogens without detectable resistance. *Nature*, **517**, 455–459.
- Little AEF, Murakami T, Mueller UG, Currie CR (2006) Defending against parasites: fungus-growing ants combine specialized behaviours and microbial symbionts to protect their fungus gardens. *Biology Letters*, **2**, 12–16.
- Lockey KH (1988) Lipids of the insect cuticle: origin, composition and function. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, **89**, 595–645.
- Madden AA, Grassetti A, Soriano J-AN, Starks PT (2013) Actinomycetes with antimicrobial activity isolated from paper wasp (Hymenoptera: Vespidae: Polistinae) nests. *Environmental Entomology*, **42**, 703–710.
- Madigan M, Martinko J (2007) *Biologie des micro-organismes*. Pearson.
- Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics (Oxford, England)*, **27**, 2957–2963.
- Marsh SE, Poulsen M, Gorosito NB *et al.* (2013) Association between *Pseudonocardia* symbionts and *Atta* leaf-cutting ants suggested by improved isolation methods.

- International Microbiology: The Official Journal of the Spanish Society for Microbiology*, **16**, 17–25.
- Marsh SE, Poulsen M, Pinto-Tomás A, Currie CR (2014) Interaction between Workers during a Short Time Window Is Required for Bacterial Symbiont Transmission in *Acromyrmex* Leaf-Cutting Ants. *PLoS ONE*, **9**, e103269.
- Martin S, Drijfhout F (2009) A Review of Ant Cuticular Hydrocarbons. *Journal of Chemical Ecology*, **35**, 1151.
- Martiny JBH, Bohannan BJM, Brown JH *et al.* (2006) Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, **4**, 102–112.
- Maschwitz U Vergleichende Untersuchungen zur Funktion der Ameisenmetathorakaldrüse. *Oecologia*, **16**, 303–310.
- Maschwitz U, Koob K, Schildknecht H (1970) Ein Beitrag zur Funktion der Metathoracaldrüse der Ameisen. *Journal of Insect Physiology*, **16**, 387–404.
- McFall-Ngai M, Hadfield MG, Bosch TCG *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences*, **110**, 3229–3236.
- McHardy IH, Goudarzi M, Tong M *et al.* (2013) Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. *Microbiome*, **1**, 17.
- Mearns-Spragg A, Bregu M, Boyd K g., Burgess J g. (1998) Cross-species induction and enhancement of antimicrobial activity produced by epibiotic bacteria from marine algae and invertebrates, after exposure to terrestrial bacteria. *Letters in Applied Microbiology*, **27**, 142–146.
- Medema MH, Fischbach MA (2015) Computational approaches to natural product discovery. *Nature Chemical Biology*, **11**, 639–648.

- Menzel F, Blüthgen N, Tolasch T *et al.* (2013) Crematoenones – a novel substance class exhibited by ants functions as appeasement signal. *Frontiers in Zoology*, **10**, 32.
- Mercier (2013) SUMATRA and SUMACLUSt: fast and exact comparison and clustering of sequences.
- Mohammadi T, Reesink HW, Vandenbroucke-Grauls CMJE, Savelkoul PHM (2005) Removal of contaminating DNA from commercial nucleic acid extraction kit reagents. *Journal of Microbiological Methods*, **61**, 285–288.
- Moran NA (2006) Symbiosis. *Current Biology*, **16**, R866–R871.
- Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. *Annual Review of Genetics*, **42**, 165–190.
- Moran NA, Sloan DB (2015) The Hologenome Concept: Helpful or Hollow? *PLOS Biol*, **13**, e1002311.
- Moreau CS, Bell CD, Vila R, Archibald SB, Pierce NE (2006) Phylogeny of the Ants: Diversification in the Age of Angiosperms. *Science*, **312**, 101–104.
- Mueller UG (2012) Symbiont recruitment versus ant-symbiont co-evolution in the attine ant-microbe symbiosis. *Current opinion in microbiology*, **15**, 269–277.
- Mueller UG, Dash D, Rabeling C, Rodrigues A (2008) Coevolution between attine ants and actinomycete bacteria: a reevaluation. *Evolution; international journal of organic evolution*, **62**, 2894–2912.
- Nannipieri P, Ascher J, Ceccherini MT *et al.* (2008) Effects of Root Exudates in Microbial Diversity and Activity in Rhizosphere Soils. In: *Molecular Mechanisms of Plant and Microbe Coexistence Soil Biology.*, pp. 339–365. Springer.
- Newman DJ, Cragg GM (2012) Natural products as sources of new drugs over the 30 years from 1981 to 2010. *Journal of Natural Products*, **75**, 311–335.

- Nordström A, O'Maille G, Qin C, Siuzdak G (2006) Nonlinear data alignment for UPLC-MS and HPLC-MS based metabolomics: quantitative analysis of endogenous and exogenous metabolites in human serum. *Analytical Chemistry*, **78**, 3289–3295.
- Nothias-Scaglia L-F, Esposito M, Costa J *et al.* (2015) Les réseaux moléculaires, une approche bio-informatique globale pour interpréter les données de spectrométrie de masse tandem. *Spectra Analyse*, 65–70.
- Oh J, Byrd AL, Deming C *et al.* (2014) Biogeography and individuality shape function in the human skin metagenome. *Nature*, **514**, 59–64.
- Oh D-C, Poulsen M, Currie CR, Clardy J (2009a) Dentigerumycin: a bacterial mediator of an ant-fungus symbiosis. *Nature Chemical Biology*, **5**, 391–393.
- Oh D-C, Poulsen M, Currie CR, Clardy J (2011) Sceliphrolactam, a Polyene Macrocyclic Lactam from a Wasp-Associated Streptomyces sp. *Organic Letters*, **13**, 752–755.
- Oh D-C, Scott JJ, Currie CR, Clardy J (2009b) Mycangimycin, a Polyene Peroxide from a Mutualist Streptomyces sp. *Organic Letters*, **11**, 633–636.
- Oliver SG, Winson MK, Kell DB, Baganz F (1998) Systematic functional analysis of the yeast genome. *Trends in Biotechnology*, **16**, 373–378.
- van Oorschot RA, Ballantyne KN, Mitchell RJ (2010) Forensic trace DNA: a review. *Investigative genetics*, **1**, 14.
- Orivel J, Leroy C (2011) The diversity and ecology of ant gardens (Hymenoptera: Formicidae; Spermatophyta: Angiospermae). *Myrmecological News*, **14**, 73–85.
- Oulhen N, Schulz BJ, Carrier TJ (2016) English translation of Heinrich Anton de Bary's 1878 speech, "Die Erscheinung der Symbiose" ("De la symbiose"). *Symbiosis*, **69**, 131–139.
- Parsons HM, Ludwig C, Günther UL, Viant MR (2007) Improved classification accuracy in 1- and 2-dimensional NMR metabolomics data using the variance stabilising generalised logarithm transformation. *BMC bioinformatics*, **8**, 234.

- Paulson JN, Stine OC, Bravo HC, Pop M (2013) Robust methods for differential abundance analysis in marker gene surveys. *Nature methods*, **10**, 1200–1202.
- Peterson J, Garges S, Giovanni M *et al.* (2009) The NIH Human Microbiome Project. *Genome Research*, **19**, 2317–2323.
- Pettit GR, Du J, Pettit RK *et al.* (2006) Antineoplastic Agents. 554. The manitoba bacterium *Streptomyces* sp. *Journal of Natural Products*, **69**, 804–806.
- Poulsen M, Oh D-C, Clardy J, Currie CR (2011) Chemical Analyses of Wasp-Associated *Streptomyces* Bacteria Reveal a Prolific Potential for Natural Products Discovery. *PLoS ONE*, **6**, e16763.
- Prescott LM, Sherwood LM, Woolverton CJ (2010) *Microbiologie*. De Boeck Supérieur.
- Prince JT, Marcotte EM (2006) Chromatographic alignment of ESI-LC-MS proteomics data sets by ordered bijective interpolated warping. *Analytical Chemistry*, **78**, 6140–6152.
- Prosser JI, Bohannan BJM, Curtis TP *et al.* (2007) The role of ecological theory in microbial ecology. *Nature Reviews Microbiology*, **5**, 384–392.
- Qiu Y, Tittiger C, Wicker-Thomas C *et al.* (2012) An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis. *Proceedings of the National Academy of Sciences*, **109**, 14858–14863.
- R Core Team (2016) R: A Language and Environment for Statistical Computing.
- R K, Sj M (2015) Evolution of Cuticular Hydrocarbons in the Hymenoptera: a Meta-Analysis., Evolution of Cuticular Hydrocarbons in the Hymenoptera: a Meta-Analysis. *Journal of chemical ecology*, *Journal of Chemical Ecology*, **41**, **41**, 871, 871–883.
- Ramadhar TR, Beemelmans C, Currie CR, Clardy J (2014) Bacterial symbionts in agricultural systems provide a strategic source for antibiotic discovery. *The Journal of Antibiotics*, **67**, 53–58.
- Rapaport F, Khanin R, Liang Y *et al.* (2013) Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biology*, **14**, 3158.

- Reber A, Chapuisat M (2011) Diversity, prevalence and virulence of fungal entomopathogens in colonies of the ant *Formica selysi*. *Insectes Sociaux*, **59**, 231–239.
- Reddy BVB, Milshteyn A, Charlop-Powers Z, Brady SF (2014) eSNaPD: A Versatile, Web-Based Bioinformatics Platform for Surveying and Mining Natural Product Biosynthetic Diversity from Metagenomes. *Chemistry & Biology*, **21**, 1023–1033.
- Riesenfeld CS, Schloss PD, Handelsman J (2004) Metagenomics: genomic analysis of microbial communities. *Annual Review of Genetics*, **38**, 525–552.
- Ripley, Venables, Billa, Bates, Douglas *et al.* (2015) Mass, package support functions and datasets for Venables and Ripley.
- Rosenberg E, Zilber-Rosenberg I (2016) Microbes Drive Evolution of Animals and Plants: the Hologenome Concept. *mBio*, **7**, e01395-15.
- Rubin BER, Sanders JG, Hampton-Marcell J *et al.* (2014) DNA extraction protocols cause differences in 16S rRNA amplicon sequencing efficiency but not in community profile composition or structure. *MicrobiologyOpen*, **3**, 910–921.
- Russell JA, Moreau CS, Goldman-Huertas B *et al.* (2009) Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. *Proceedings of the National Academy of Sciences*, **106**, 21236–21241.
- Salem H, Florez L, Gerardo N, Kaltenpoth M (2015) An out-of-body experience: the extracellular dimension for the transmission of mutualistic bacteria in insects. *Proceedings of the Royal Society B: Biological Sciences*, **282**.
- Salter SJ, Cox MJ, Turek EM *et al.* (2014) Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC biology*, **12**, 87.
- Sambrook J, Fritschi EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, New York.
- Scheuring I, Yu DW (2012) How to assemble a beneficial microbiome in three easy steps. *Ecology Letters*, **15**, 1300–1307.

- Schnell IB, Bohmann K, Gilbert MTP (2015) Tag jumps illuminated – reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*, **15**, 1289–1303.
- Schoenian I, Spiteller M, Ghaste M *et al.* (2011) Chemical basis of the synergism and antagonism in microbial communities in the nests of leaf-cutting ants. *Proceedings of the National Academy of Sciences*, **108**, 1955–1960.
- Schoeters E, Billen J The intramandibular gland, a novel exocrine structure in ants (Insecta, Hymenoptera). *Zoomorphology*, **114**, 125–131.
- Schoeters, E., Billen, J. (1991) Morphological and ultrastructural study of the metapleural gland in *Diacamma* (Hymenoptera, Formicidae). *Proceedings Leuven*, 239–247.
- Scott JJ, Oh D-C, Yuceer MC *et al.* (2008) Bacterial Protection of Beetle-Fungus Mutualism. *Science*, **322**, 63–63.
- Seipke RF, Barke J, Brearley C *et al.* (2011) A Single *Streptomyces* Symbiont Makes Multiple Antifungals to Support the Fungus Farming Ant *Acromyrmex octospinosus*. *PLOS ONE*, **6**, e22028.
- Seipke RF, Barke J, Heavens D, Yu DW, Hutchings MI (2013) Analysis of the bacterial communities associated with two ant–plant symbioses. *MicrobiologyOpen*, **2**, 276–283.
- Sen R, Ishak HD, Estrada D *et al.* (2009) Generalized antifungal activity and 454-screening of *Pseudonocardia* and *Amycolatopsis* bacteria in nests of fungus-growing ants. *Proceedings of the National Academy of Sciences*, **106**, 17805–17810.
- Shao Z, Luo Y, Zhao H (2011) Rapid Characterization and Engineering of Natural Product Biosynthetic Pathways via DNA Assembler. *Molecular bioSystems*, **7**, 1056–1059.
- Sheffer E, Batterman SA, Levin SA, Hedin LO (2015) Biome-scale nitrogen fixation strategies selected by climatic constraints on nitrogen cycle. *Nature Plants*, **1**, 15182.

- Shinzato C, Inoue M, Kusakabe M (2014) A Snapshot of a Coral “Holobiont”: A Transcriptome Assembly of the Scleractinian Coral, *Porites*, Captures a Wide Variety of Genes from Both the Host and Symbiotic *Zooxanthellae*. *PLOS ONE*, **9**, e85182.
- Smith CA, Want EJ, O’Maille G, Abagyan R, Siuzdak G (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical Chemistry*, **78**, 779–787.
- Stasulli NM, Shank EA (2016) Profiling the metabolic signals involved in chemical communication between microbes using imaging mass spectrometry. *FEMS Microbiology Reviews*, **40**, 807–813.
- Sturgis, S.J., Gordon, D.M. (2012) Nestmate recognition in ants (Hymenoptera: Formicidae): a review. *Myrmecological News*, **16**, 101–110.
- Suen G, Scott JJ, Aylward FO *et al.* (2010) An Insect Herbivore Microbiome with High Plant Biomass-Degrading Capacity. *PLOS Genet*, **6**, e1001129.
- Sylvain F-É, Cheaib B, Llewellyn M *et al.* (2016) pH drop impacts differentially skin and gut microbiota of the Amazonian fish tambaqui (*Colossoma macropomum*). *Scientific Reports*, **6**, 32032.
- Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, **21**, 2045–2050.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, **30**, 2725–2729.
- Tautenhahn R, Böttcher C, Neumann S (2008) Highly sensitive feature detection for high resolution LC/MS. *BMC bioinformatics*, **9**, 504.

- Theis KR, Dheilly NM, Klassen JL *et al.* (2016) Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes. *mSystems*, **1**, e00028-16.
- Thomsen PF, Willerslev E (2015) Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, **183**, 4–18.
- Tindall BJ, Rosselló-Móra R, Busse H-J, Ludwig W, Kämpfer P (2010) Notes on the characterization of prokaryote strains for taxonomic purposes. *International Journal of Systematic and Evolutionary Microbiology*, **60**, 249–266.
- Tokovenko B, Rebets Y, Luzhetskyy A (2016) Automating Assessment of the Undiscovered Biosynthetic Potential of Actinobacteria. *bioRxiv*, 1–27.
- Turnbaugh PJ, Gordon JI (2008) An Invitation to the Marriage of Metagenomics and Metabolomics. *Cell*, **134**, 708–713.
- Turnbaugh PJ, Ley RE, Hamady M *et al.* (2007) The human microbiome project. *Nature*, **449**, 804–810.
- Ule E (1901) Ameisengärten im Amazonasgebiet. *Beiblatt Bot. Jahrb.*, 45–52.
- Ushio M, Yamasaki E, Takasu H *et al.* (2015) Microbial communities on flower surfaces act as signatures of pollinator visitation. *Scientific Reports*, **5**, 8695.
- Vallet-Gely I, Lemaitre B, Bocard F (2008) Bacterial strategies to overcome insect defences. *Nature Reviews Microbiology*, **6**, 302–313.
- Valli S, Suvathi SS, Aysha OS *et al.* (2012) Antimicrobial potential of Actinomycetes species isolated from marine environment. *Asian Pacific Journal of Tropical Biomedicine*, **2**, 469–473.
- Velho-Pereira S, Kamat NM (2011) Antimicrobial Screening of Actinobacteria using a Modified Cross-Streak Method. *Indian Journal of Pharmaceutical Sciences*, **73**, 223.
- Vincent C, Manges AR (2015) Antimicrobial Use, Human Gut Microbiota and *Clostridium difficile* Colonization and Infection. *Antibiotics (Basel, Switzerland)*, **4**, 230–253.

- Vo A-TE, Jedlicka JA (2014) Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for faecal and swab samples. *Molecular Ecology Resources*, **14**, 1183–1197.
- Wang M, Carver JJ, Phelan VV *et al.* (2016) Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*, **34**, 828–837.
- Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.
- Wang Y, Mueller UG, Clardy J (1999) Antifungal Diketopiperazines from Symbiotic Fungus of Fungus-Growing Ant *Cyphomyrmex minutus*. *Journal of Chemical Ecology*, **25**, 935–941.
- Ward DM, Cohan FM, Bhaya D *et al.* (2008) Genomics, environmental genomics and the issue of microbial species. *Heredity*, **100**, 207–219.
- Wesolowska-Andersen A, Bahl MI, Carvalho V *et al.* (2014) Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. *Microbiome*, **2**, 19.
- Wetterer JK (1999) The ecology and evolution of worker size-distribution in leaf-cutting ants (Hymenoptera: Formicidae). *SOCIOBIOLOGY*, **34**, 119–144.
- Wheeler WM (1921) A new case of parabiosis and the “ant gardens” of British Guiana. *Ecology*, 89–103.
- Wiegand I, Hilpert K, Hancock REW (2008) Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nature Protocols*, **3**, 163–175.
- van Wilgenburg E, Symonds MRE, Elgar MA (2011) Evolution of cuticular hydrocarbon diversity in ants. *Journal of Evolutionary Biology*, **24**, 1188–1198.

- Williston EH, Zia-Walrath P, Youmans GP (1947) Plate Methods for Testing Antibiotic Activity of Actinomycetes against Virulent Human Type Tubercle Bacilli. *Journal of Bacteriology*, **54**, 563–568.
- Willner D, Daly J, Whiley D *et al.* (2012) Comparison of DNA Extraction Methods for Microbial Community Profiling with an Application to Pediatric Bronchoalveolar Lavage Samples. *PLOS ONE*, **7**, e34605.
- Wilson EO (1980) Caste and division of labor in leaf-cutter ants (Hymenoptera: Formicidae: Atta). *Behavioral Ecology and Sociobiology*, **7**, 157–165.
- Xia J, Wishart DS (2011) Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nature Protocols*, **6**, 743–760.
- Yang JY, Sanchez LM, Rath CM *et al.* (2013) Molecular networking as a dereplication strategy. *Journal of Natural Products*, **76**, 1686–1699.
- Yek SH, Mueller UG (2011) The metapleural gland of ants. *Biological Reviews*, **86**, 774–791.
- Yoshimasa H, Hisao N, Kiyoyuki Y, Kunio O, Takayuki N (1961) The structure of aureothin, a nitro compound obtained from *Streptomyces thioluteus*. *Tetrahedron*, **14**, 252–274.
- Zinger L, Chave J, Coissac E *et al.* (2016) Extracellular DNA extraction is a fast, cheap and reliable alternative for multi-taxa surveys based on soil DNA. *Soil Biology and Biochemistry*, **96**, 16–19.
- Zipperer A, Konnerth MC, Laux C *et al.* (2016) Human commensals producing a novel antibiotic impair pathogen colonization. *Nature*, **535**, 511–516.
- Zucchi TD, Guidolin AS, Cônsoli FL (2011) Isolation and characterization of actinobacteria ectosymbionts from *Acromyrmex subterraneus brunneus* (Hymenoptera, Formicidae). *Microbiological Research*, **166**, 68–76.