Université de Montréal

Étude de la biodiversité microbienne associée aux champignons mycorhiziens arbusculaires dans des sites hautement contaminés par des hydrocarbures pétroliers

par
Bachir Iffis

Département de Sciences Biologiques
Institut de Recherche en Biologie Végétale - Centre sur la Biodiversité
Faculté des Arts et des Sciences

Thèse présentée à la Faculté des Arts et des Sciences
en vue de l'obtention du grade de Philosophiae Doctor (PhD)
en Sciences Biologiques

Juillet 2016

Université de Montréal
Faculté des Études Supérieures et Postdoctorales

Cette thèse intitulée:

Étude de la biodiversité microbienne associée aux champignons mycorhiziens arbusculaires dans des sites hautement contaminés par des hydrocarbures pétroliers

Présentée par:
Bachir Iffis

a été évaluée par un jury composé des personnes suivantes:

Dr. Marc St-Arnaud, président-rapporteur
Dr. Mohamed Hijri, directeur de recherche
Dr. Jesse Shapiro, membre du jury
Dr. Philippe Constant, examinateur externe

Résumé

Les champignons mycorhiziens à arbuscules (CMA) forment un groupe de champignons qui appartient à l'embranchement des Glomeromycètes (Glomeromycota). Les CMA forment des associations symbiotiques, connus sous le nom des mycorhizes à arbuscules avec plus de 80 % des plantes vasculaires terrestres. Une fois que les CMA colonisent les racines de plantes, ils améliorent leurs apports nutritionnels, notamment le phosphore et l'azote, et protègent les plantes contre les différents pathogènes du sol. En contrepartie, les plantes offrent un habitat et les ressources de carbone nécessaires pour le développement et la reproduction des CMA. Des études plus récentes ont démontré que les CMA peuvent aussi jouer des rôles clés dans la phytoremédiation des sols contaminés par les hydrocarbures pétroliers (HP) et les éléments traces métaliques. Toutefois, dans les écosystèmes naturels, les CMA établissent des associations tripartites avec les plantes hôtes et les microorganismes (bactéries et champignons) qui vivent dans la rhizosphère, l'endosphère (à l'intérieur des racines) et la mycosphère (sur la surface des mycéliums des CMA), dont certains d'entre eux jouent un rôle dans la translocation, l'immobilisation et/ou la dégradation des polluants organiques et inorganiques présents dans le sol. Par conséquent, la diversité des CMA et celle des microorganismes qui leur sont associés sont influencées par la concentration et la composition des polluants présents dans le sol, et aussi par les différents exsudats sécrétés par les trois partenaires (CMA, bactéries et les racines de plantes). Cependant, la diversité des CMA et celle des microorganismes qui leur sont associés demeure très peu connue dans les sols contaminés. Les interactions entre les CMA et ces microorganismes sont aussi méconnus aussi bien dans les aires naturelles que contaminées.

Dans ce contexte, les objectifs de ma thèse sont: i) étudier la diversité des CMA et les microorganismes qui leur sont associés dans des sols contaminés par les HP, ii) étudier la variation de la diversité des CMA ainsi que celle des microorganismes qui leur sont associés par rapport au niveau de concentration en HP et aux espèces de plantes hôtes, iii) étudier les correlations (covariations) entre les CMA et les microorganismes qui leur sont associés et iv) comparer les communautés microbiennes trouvées dans les racines et sols contaminés par les HP avec celles trouvées en association avec les CMA.

Pour ce faire, des spores et/ou des propagules de CMA ont été extraites à partir des racines et des sols de l'environnement racinaire de trois espèces de plantes qui poussaient spontanément dans trois bassins de décantation d'une ancienne raffinerie de pétrole située dans la Rive-Sud du fleuve St-Laurent, près de Montréal. Les spores et les propagules collectées, ainsi que des échantillons du sol et des racines ont été soumis à des techniques de PCR (nous avons ciblés les genes 16S de l'ARNr pour bactéries, les genes 18S de l'ARNr pour CMA et les régions ITS pour les autres champignons), de clonage, de séquençage de Sanger ou de séquençage à haut débit. Ensuite, des analyses bio-informatiques et statistiques ont été réalisées afin d'évaluer les effets des paramètres biotiques et abiotiques sur les communautés des CMA et les microorganismes qui leur sont associés.

Mes résultats ont montré une diversité importante de bactéries et de champignons en association avec les spores et les propagules des CMA. De plus, la communauté microbienne associée aux spores des CMA a été significativement affectée par l'affiliation taxonomique des plantes hôtes et les niveaux de concentration en HP. D'autre part, les corrélations positives ou négatives qui ont été observées entre certaines espèces de CMA et microorganismes suggérèrent qu’en plus des effets de la concentration en HP et l'identité des plantes hôtes, les

CMA peuvent aussi affecter la structure des communautés microbiennes qui vivent sur leurs spores et mycéliums. La comparaison entre les communautés microbiennes identifiées en association avec les spores et celles identifiées dans les racines montre que les communautés microbiennes recrutées par les CMA sont différentes de celles retrouvées dans les sols et les racines.

En conclusion, mon projet de doctorat apporte de nouvelles connaissances importantes sur la diversité des CMA dans un environnement extrêmement pollué par les HP, et démontre que les interactions entre les CMA et les microorganismes qui leur sont associés sont plus compliquées que ce qu'on croyait précédemment. Par conséquent, d'autres travaux de recherche sont recommandés, dans le futur, afin de comprendre les processus de recrutement des microorganismes par les CMA dans les différents environnements.

Mots-clés : champignons mycorhiziens à arbuscules (CMA), hydrocarbures pétroliers (HP), microorganismes associés aux CMA, bactéries, champignons, séquençage à haut débit, clonage, séquençage de Sanger.

Abstract

Arbuscular mycorrhizal fungi (AMF) are an important soil fungal group that belongs to the phylum Glomeromycota. AMF form symbiosic associations known as arbuscular mycorrhiza with more than 80% of vascular plants on earth. Once AMF colonize plant roots, they promote nutrient uptake, in particular phosphorus and nitrogen, and protect plants against soil-borne pathogens. In turn, plants provide AMF with carbon resources and habitat. Furthermore, more recent studies demonstrated that AMF may also play key roles in phytoremediation of soils contaminated with petroleum hydrocarbon pollutants (PHP) and trace elements. Though, in natural ecosystems, AMF undergo tripartite associations with host plants and micoorganisms (Bacteria and Fungi) living in rhizosphere (the narrow region of soil surounding the plant roots), endosphere (inside roots) and mycosphere (on the surface AMF mycelia), which some of them play a key role on translocation, immobilization and/or degradation of organic and inorganic pollutants. Consequently, the diversity and community structures of AMF and their associated microorganisms are influenced by the composition and concentration of pollutants and exudates released by the three partners (AMF, bacteria and plant roots). However, little is known about the diversity of AMF and their associated microorganisms in polluted soils and the interaction between AMF and these microorganisms remains poorly understood both in natural and contaminated areas.

In this context, the objectives of my thesis were to: i) study the diversity of AMF and their associated microorganisms in PHP contaminated soils, ii) study the variation in diversity and community structures of AMF and their associated microorganisms across plant species identity and PHP concentrations, iii) study the correlations (covariations) between AMF

species and their associated microorganisms and iv) compare microbial community structures of PHP contaminated soils and roots with those associated with AMF spores in order to determine if the microbial communities shaped on the surface of AMF spores and mycelia are different from those identified in soil and roots.

To do so, AMF spores and/or their intraradical propagules were harvested from rhizospheric soil and roots of three plant species growing spontaneously in three distinct waste decantation basins of a former petrochemical plant located on the south shore of the St-Lawrence River, near Montreal. The harvested spores and propagules, as well as samples of soils and roots were subjected to PCR (we target 16S rRNA genes for bacteria, 18S rRNA genes for AMF and ITS regions for the other fungi), cloning, Sanger sequencing or 454 high throughput sequencing. Then, bioinformatics and statistics were performed to evaluate the effects of biotic and abiotic driving forces on AMF and their associated microbial communities.

My results showed high fungal and bacterial diversity associated with AMF spores and propagules in PHP contaminated soils. I also observed that the microbial community structures associated with AMF spores were significantly affected by plant species identity and PHP concentrations. Furthermore, I observed positive and negative correlations between some AMF species and some AMF-associated microorganisms, suggesting that in addition to PHP concentrations and plant species identity, AMF species may also play a key role in shaping the microbial community surrounding their spores. Comparisons between the AMF spore-associated microbiome and the whole microbiome found in rhizospheric soil and roots showed that AMF spores recruit a microbiome differing from those found in the surrounding soil and roots.

Overall, my PhD project brings a new level of knowledge on AMF diversity on extremely polluted environment and demonstrates that interaction of AMF and their associated microbes is much complex that we though previously. Further investigations are needed to better understand how AMF select and reward their associated microbes in different environments.

Keywords : Arbuscular Mycorrhizal Fungi (AMF), petroleum hydrocarbons pollutants (PHP), AMF-associated microorganisms, bacteria, fungi, 454 high throughput sequencing, cloning, Sanger sequencing.

Table des matières

Liste des tableaux

Liste des figures

Liste des abréviations et des sigles

ANOVA	analysis de la variance
ADN	acide désoxyribonucléique
AMF	arbuscular mycorrhizal fungi
ARN	acide ribonucléique
ARNr	ARN ribosomal
BLASTN	basic local alignment search tool nucleotide
bp	base pairs
cm	centimètre
CMA	champignons mycorhiziens à arbuscule
CoIA	co-inertia analysis
Cu	cuivre
DNA	desoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
dNTP	deoxynucleotide triphosphate
h	heure
HAP	hydrocarbures aromatiques polycycliques
HC	high contamination,
HP	hydrocarbures pétroliers
ITS	internal transcribed spacer
kg	kilogramme
KOH	hydroxyde de potassium
L	litre
LC	low contamination.
MC	moderate contamination,
MDDELCC	ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques
mg	milligramme
$MgCl_2$	chlorure de magnésium

MHB	mycorrhiza helper bacteria
MID	unique multiplex identifier
min	minute
mL	millilitre
mM	millimolaire
N	north
NCBI	National center for biotechnology information
ng	nanogramme
NMDS	non-metric multidimensional scaling
NSERC	Natural Science and Engineering Research Council of Canada
OTU	operational taxonomic unit
PAH	polycyclic aromatic hydrocarbons
PAH	polycyclic aromatic hydrocarbons
Pb	plomb
PCoA	principal coordinates analysis
PCR	polymerase chain reaction
PDA	potato dextrose agar
PERMANOVA	Permutational Multivariate Analysis of Variance
PHP	petroleum hydrocarbon pollutants
PSB	phosphate-solubilizing bacteria
rDNA	ribosomal DNA
RNA	ribonucleic acid
rRNA	ribosomal RNA
s	seconde
SEM	scanning electron microscopy
sp.	species (espèce)
TSA	trypto-caséine soja
U	unit
UV	ultraviolet
W	west
WGA	whole-genome amplification

Zn	zinc
α	alpha
β	beta
γ	gamma
µL	microlitre
µM	micromolaire
µm	micromètre
3D	trois dimensions

«Cette thèse de doctorat est dédiée à mes chers parents qui vivent loin de moi (Algérie). Leurs prières et leurs bénédictions m'ont toujours suivi dans mon parcours. Je suis certain qu'il ne se passe pas un jour sans qu'ils pensent à moi. Que Dieu vous protège et vous garde pour moi.

Aussi à ma femme, Assa Djazia Kichou, qui a partagé avec moi tous les moments difficiles que j'ai rencontrés durant cette thèse de doctorat. Tous mes profonds respects et reconnaissances pour tes sacrifices, ta patience, ton courage et ta bienveillance.»

Remerciements

« La terre...Combien sommes-nous à comprendre cette glèbe silencieuse que nous foulons toute notre vie? Pourtant…C'est elle qui nous nourrit, elle à qui nous devons la vie et devrons irrévocablement la survie »
-Pierre Rabhi

J'ai intégré le laboratoire de recherche de l'IRBV pour un stage de M2 de l'Université de Dijon en France pour une durée de 6 mois et je n'aurais jamais cru que j'allais initier un doctorat et m'installer au Canada. Mon projet de thèse est un aboutissement de mon parcours académique et professionnel. Il y a eu différents obstacles qui m'ont permis d'en sortir plus fort et gagnant. De ce fait, ce projet de thèse n'aurait pas été possible sans le soutien de certaines personnes.

Mes premiers remerciements vont d'abord à Dr. Mohamed Hijri, directeur de thèse, de m'avoir accepté dans son laboratoire de recherche et de m'avoir accompagné tout au long de ma formation. Je remercie également Dr. Marc-St-Arnaud qui a été présent aux moments opportuns et qui a été d'une grande aide dans les moments les plus dures. Il m'a apporté de judicieux conseils et suggestions.

Je remercie tous les membres du jury d'avoir accepté l'évaluation de ma thèse. Merci à Dr. Jesse Shapiro, professeur adjoint à l'Université de Montréal. Merci à Dr. Philippe Constant, chercheur du Centre INRS-Institut Armand-Frappier. Merci à Dr Marc St-Arnaud, botaniste chercheur du Jardin botanique de Montréal et professeur associé de l'Université de Montréal.

Je remercie particulièrement les membres du laboratoire Hijri. Entre autres, un merci à Maryam Nadimi et Denis Beaudet de m'avoir permis de collaborer sur leurs projets, à

Stéphanie Berthiaume, Charlotte Marchand, Laurence Daubois, Alice Roy-Bolduc, Saad El-Din Hassan, Youssef Ismail, Abdelghani El Yassimi, Rim Ben Haj Sassi, Guillaume Bourdel, Manuel Labridy, Dimitri Dagher et MengXuan Kong (Cathy) d'avoir partagé le laboratoire. Un grand merci aux membres du personnel de l'IRBV.

Mes remerciements vont également à Dr. Franck Stefani, à Stéphane Daigle, à Dr. Rim Klabi, à Louise Pelletier, à Dr. Karen Fisher-Favret et à Dr. Terrence Bell pour le soutien technique, pour le support en statistique et pour l'aide dans l'analyse des séquences.

Ce projet de recherche a été financé par GenoRem qui soutient une approche de phytoremédiation innovatrice pour décontaminer les sols pollués tout en respectant les principes du développement durable. J'ai eu l'occasion de participer à la 7e Conférence internationale sur les mycorhizes à New Delhi en Inde en 2013 grâce, entre autres, à la bourse de déplacement du Centre de la science de la biodiversité du Québec (CSBQ). Je remercie également la Faculté des Études Supérieures et Post-doctorales (FESP) de l'Université de Montréal pour l'obtention de la bourse de fin d'études doctorales.

Je remercie du fond du cœur mes parents pour leur soutien inconditionnel et à toute ma famille pour leur encouragement malgré la distance qui nous sépare. Un merci à mes beaux-parents pour leur soutien moral et gastronomique.

Un énorme remerciement à mon épouse qui a été présente tout au long de ce parcours, qui a su m'encourager et m'écouter dans les moments les plus difficiles, qui a su me soutenir pour aller jusqu'au bout de cette thèse. Merci à ma chère épouse Assa Djazia Kichou.

Merci à tous mes amis intercontinentaux et à mon équipe de jogging.

Chapitre 1 - Introduction générale

1.1. Les défis de la remédiation des hydrocarbures pétroliers

Les hydrocarbures pétroliers (HP) et leurs produits dérivés sont largement utilisés dans la vie quotidienne humaine, en particulier dans la production d'énergie, la construction, le transport et la production industrielle. Toutefois, l'exploitation intensive des ressources pétrolières, notamment les hydrocarbures aromatiques polycycliques (HAP), peut générer de sérieux problèmes environnementaux qui peuvent avoir de graves conséquences sur le sol, l'air, l'eau, la faune et la flore (Boffetta *et al.*, 1997; Samanta *et al.*, 2002; Poirier, 2004; McAloose & Newton, 2009). Par exemple, la catastrophe de déversement de pétrole brut qui est survenue le 6 juillet 2013 au Lac-Mégantic (Québec) a eu des conséquences très graves sur l'environnement. Selon la presse locale, environ 100 000 litres de pétrole brut ont été déversés sur une surface de 80 km dans la rivière Chaudière, tandis que les concentrations en HAP et arsenic ont été 394 444 et 28 fois, respectivement, supérieures aux normes admissibles par le ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques (MDDELCC)[1]. La décontamination des sols pollués par les HP est difficile, longue et très coûteuse. L'excavation, le transport et le déversement du sol contaminé dans des sites d'enfouissement sécurisés demeure l'approche conventionnelle la plus utilisée. Par contre, cette méthode est non seulement coûteuse, mais elle ne résout pas réellement le problème de décontamination du sol, car celle-ci ne fait que déplacer les contaminants d'un site à un autre. D'autre part, le caractère hydrophobe élevé des HP et de leur capacité à

[1] Les données ont été obtenus à partir du journal "Le Devoir" (http://www.ledevoir.com) publié le 14 août 2013 et du journal "L'actualité" (http://www.lactualite.com) publié le 11 juillet 2013.

s'adsorber aux matières organiques du sol réduisent considérablement leur hydro-solubilité et biodisponibilité, par conséquent cela réduit considérablement l'efficacité des procédures chimiques de décontamination. En outre, les réactifs chimiques utilisés dans la dépollution des HP tels que l'ozone, le peroxyde d'hydrogène, le permanganate et le persulfate peuvent interagir avec la matière organique du sol et de modifier radicalement les critères physico-chimiques et microbiologiques des sols (Mahanty *et al.*, 2011). Pour faire face à ces problèmes, la phytoremédiation qui consiste à utiliser des plantes pour la décontamination des sols (Peuke & Rennenberg, 2005) est l'alternative la plus écologique, la plus harmonieuse avec le paysage, et la moins coûteuse. Cependant, dans les milieux naturels, les communautés bactériennes et fongiques de la rhizosphère jouent un rôle primordial dans les processus de la phytoremédiation. Ainsi, comprendre ces processus nécessite une connaissance approfondie des communautés microbiennes présentes dans les sols entourant les racines des plantes et vivant à l'intérieur ce celles-ci.

1.2. Généralités sur les champignons mycorhiziens à arbuscules

Les champignons mycorhiziens à arbuscules (CMA) appartiennent à l'embranchement de Glomeromycota (Schüβler *et al.*, 2001). Selon les révisions taxonomiques réalisées par Oehl *et al.* (2011), les Glomérомycètes sont représentés par environ 230 espèces réparties en trois classes, cinq ordres, quatorze familles et vingt-neuf genres. L'ensemble des CMA sont des organismes telluriques (vivent dans les sols) possédant des hyphes coénocytiques et des spores multi-nucléées dont le diamètre varie en moyenne entre 50 µm et 500 µm (Fortin *et al.*, 2008; Smith & Read, 2008; Marleau *et al.*, 2011) (Figure 1.1-A). Cependant, certaines espèces telle que *Gigaspora gigantea*, le diamètre peut dépasser 750 µm. Quand les CMA colonisent les racines des plantes hôtes, les hyphes extra-racinaires traversent les parois des cellules

corticales et forment des interfaces des échanges bidirectionnels des nutriments avec le cytoplasme appelées arbuscules, du latin *arbusculum* qui signifie arbuste. En plus des arbuscules, chez certaines espèces de CMA, les parties des hyphes qui ne traversent pas les cellules corticales peuvent se différencier en structures sphériques multi-nucléées connues sous le nom de vésicules. Selon Smith and Read (2008), les vésicules jouent un rôle d'organe de stockage des réserves nutritionnelles pour les CMA. Certaines espèces de CMA telle que *Rhizophagus intraradices* forment aussi des spores intra-racinaires qui ressemblent aux vésicules. Puisqu'il est difficile de distinguer entre les vésicules et les spores intra-racinaires, le terme propagules intra-racinaires est utilisé pour désigner l'ensemble des vésicules et spores intra-racinaires (Figure 1.1-B). Dans le chapitre 2 de ce rapport de thèse, ces structures intra-racinaires ont fait l'objet d'études de la biodiversité des microorganismes associés aux CMA.

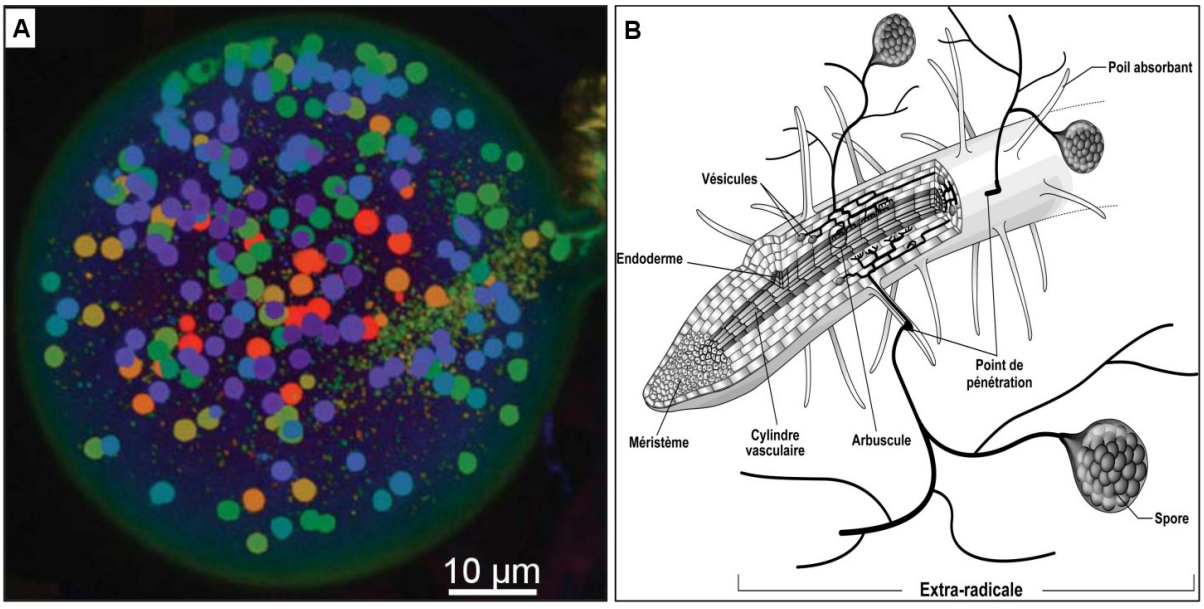


Figure 1.1: Morphologie des CMA. (A) images de microscopie confocale montrant une spore de CMA qui contient une centaine de noyaux à l'intérieur de son cytoplasme (Marleau *et al.*, 2011), (B) un schéma des différentes structures extra et intra racinaires des CMA (Fortin *et al.*, 2008).

Quant à leur mode de vie, les CMA sont des biotrophes obligatoires formant des associations symbiotiques avec plus de 80 % des plantes vasculaires, ce qui représente approximativement 250 000 espèces de plantes (Smith & Read, 2008). L'association des CMA avec les plantes est considérée comme étant l'une des premières formes de symbiose entre les microorganismes et les plantes terrestres. D'après les études paléontologiques sur les fossiles dévoniens de *Rhynia* spp, cette relation étroite entre les CMA et les plantes date d'environ 450 millions d'années (Simon *et al.*, 1993; Taylor *et al.*, 1995; Redecker *et al.*, 2000).

Les rôles bénéfiques des CMA ont fait l'objet de plusieurs travaux de recherche qui ont clairement démontré en laboratoire et en champs les impacts positifs de ces champignons sur la croissance des plantes, l'augmentation des rendements en agriculture et la protection des plantes contre les stresses biotiques (e.g. les pathogènes) et abiotiques (e.g. pollution, sécheresse et salinité). Dans le domaine de l'agriculture, ces champignons sont considérés comme étant des bio-fertilisants par excellence. Lorsque les racines de plantes sont colonisées par les CMA, ces derniers développent un réseau d'hyphes extra-racinaire, qui s'étendent au-delà de la zone exploitée par les racines en permettant ainsi une meilleure exploitation de l'eau et les sels minéraux qui se trouvent en dehors de la zone rhizosphérique (Cho *et al.*, 2009; Hu *et al.*, 2010; Abdel Latef & Chaoxing, 2011; Karagiannidis *et al.*, 2011). De plus, plusieurs études ont démontré que l'association mycorhizienne arbusculaire confère aux plantes une résistance accrue contre certains phyto-pathogènes du sol (St-Arnaud & Vujanovic, 2007; Ismail *et al.*, 2011; Ismail *et al.*, 2013; Schouteden *et al.*, 2015; Song *et al.*, 2015).

Dans le domaine de l'environnement, les CMA peuvent jouer des rôles clés dans la phytoremédiation des sols contaminés par les hydrocarbures pétroliers (HP) et les éléments-traces métalliques (Liu & Dalpé, 2009; Gao *et al.*, 2011b; Hassan *et al.*, 2013; Cabral *et al.*,

2015). Par exemple, Liu and Dalpé (2009) ont conduit une expérience en serre et ils ont observé que, après 12 jours, les concentrations en anthracène et phénanthrène peuvent être réduites jusqu'a 80 % dans les pots cultivés avec des poireaux mycorhizés. Par ailleurs, une expérience similaire a été réalisée par Hassan *et al.* (2013) sur des sols contaminés par les éléments-traces métalliques (zinc, cadmium et cuivre) et ils ont constaté que l'inoculation des tournesols avec l'espèce *Rhizophagus irregularis* augmentait significativement la capacité de phyto-extraction de ces contaminants par rapport aux sols cultivés par des tournesols non-mycorhizés.

1.3. Les microorganismes associés aux CMA

Dans un environnement aussi diversifié que complexe comme le sol, l'association symbiotique entre les CMA et les plantes ne dépend pas uniquement de ces deux partenaires. Plusieurs autres facteurs chimiques, biologiques et climatiques peuvent avoir un grand impact sur cette association mycorhizienne (Santos-González *et al.*, 2007; Posada *et al.*, 2008).

La biodiversité bactérienne du sol est considérée parmi les facteurs majeurs qui peuvent influencer la symbiose mycorhizienne. En plus de leur association avec les racines de plantes, les CMA sont aussi capables d'établir d'autres associations symbiotiques ou parasitiques avec les populations bactériennes et fongiques du sol (Tarkka & Frey-Klett, 2008; Bonfante & Anca, 2009; Lecomte *et al.*, 2011; Sundram *et al.*, 2011; Yasmeen *et al.*, 2012). Les associations tripartites entre les CMA, les plantes et les autres microorganismes rendent l'étude des différentes interactions entre les trois partenaires très difficile à réaliser. Parmi les populations microbiennes associées aux CMA, des bactéries appartenant aux taxons des *Firmicutes*, *Bacillus*, *Actinobacteria*, *Flexibacter*, *Cyanobacteria*, *Alpha-* et *Beta-*

proteobacteria sont capables de former des biofilms intimement associés à la surface des spores et mycélia des CMA (Roesti *et al.*, 2005; Long *et al.*, 2008; Scheublin *et al.*, 2010; Lecomte *et al.*, 2011) (Figure 1. 2-A et B). Des champignons appartenant aux embranchements *Ascomycota* et *Chytridiomycota* ont été aussi identifiés à la surface des spores CMA (Sylvia & Schenck, 1983; Tzean *et al.*, 1983; Paulitz & Menge, 1984; Hijri *et al.*, 2002). En plus de ces communautés microbiennes qu'on trouve en surface, plusieurs études ont démontré que les CMA hébergent des bactéries à l'intérieur de leurs spores et mycélia (MacDonald *et al.*, 1982; Bianciotto *et al.*, 1996; Bianciotto *et al.*, 2000) (Figure 1. 2-C et D). En utilisant les techniques de la microscopie confocale, Bianciotto *et al.* (1996) ont trouvé qu'une spore de CMA peut contenir jusqu'à 250 000 endo-bactéries à l'intérieur de son cytoplasme. Malgré que les premières observations de ces bactéries dataient des années 1970 (Mosse, 1970), on connaît peu de choses sur la biodiversité, le rôle, l'origine et les mécanismes d'infection et de transmission de ces endo-bactéries. En raison de leur mode de vie présumé biotrophe obligatoire (Jargeat *et al.*, 2004), l'identification de ces endo-bactéries est généralement limitée à des techniques d'observation microscopique et d'amplification des régions spécifiques des séquences 16S de l'ARN ribosomique (Bianciotto *et al.*, 1996; Bianciotto *et al.*, 2003; Naumann *et al.*, 2010). Récemment, les progrès qu'ont connus les techniques de génomique, protéomique et transcriptomique ont permis aux chercheurs d'avoir plus d'informations sur les interactions et les voies métaboliques de ces endo-bactéries (Ghignone *et al.*, 2012; Salvioli *et al.*, 2016; Vannini *et al.*, 2016). En utilisant les techniques de clonage et de séquençage à haut débit, Ghignone *et al.* (2012) ont réussi à séquencer le génome complet d'une endo-bactérie typique, *Candidatus Glomeribacter gigasporarum*, des *Gigasporaceae* (une famille des CMA). Dans une autre étude plus récente, Salvioli *et al.*

(2016) ont étudié le transcriptome des *Candidatus Glomeribacter gigasporarum* et ils ont démontré que cette endo-bactérie est une endo-symbionte obligatoire des *Gigasporaceae* capable d'améliorer l'état physiologique des CMA et influence leurs interactions avec les racines de plantes.

D'autre part, les mécanismes de recrutement des microorganismes vivant à la surface des spores des CMA demeurent aussi méconnus. Cependant, des hypothèses suggèrent que les CMA adoptent des stratégies similaires à celles des racines de plantes en sécrétant des sucres, des acides organiques, des acides aminés et d'autres molécules de signalisation par lesquelles ils sélectionnent la flore microbienne de leur mycosphère (Roesti *et al.*, 2005; Bharadwaj *et al.*, 2011). Quant à leurs rôles, il a été démontré qu'un groupe de bactéries nommé *mycorrhiza helper bacteria* (MHB), incluant des bactéries solubilisatrices du phosphore et fixatrice d'azote, est capable d'établir des associations symbiotiques avec les CMA (Garbaye, 1994; Tarkka & Frey-Klett, 2008; Bonfante & Anca, 2009). En échange d'un habitat et des ressources de carbone, les MHB permettent aux CMA d'augmenter leurs capacités d'absorption et de translocation des sels minéraux, en sécrétant des enzymes et des acides organiques capables de solubiliser le phosphore et de fixer l'azote atmosphérique (Bonfante & Anca, 2009; Miransari, 2011; Taktek *et al.*, 2015). De plus, il a été démontré que les MHB améliorent les taux de mycorhizations des racines de plantes par la production des composés de signalisation capables de stimuler les dialogues chimiques entre les plantes et les CMA (Garbaye, 1994; Marschner & Timonen, 2006; Tarkka & Frey-Klett, 2008; Bonfante & Anca, 2009).

Dans le domaine de la phytoremédiation, certaines études ont clairement montré que la co-inoculation des racines de plantes par les CMA et certaines espèces de bactéries peut avoir

un impact positif sur la décontamination des sols pollués et confère aux plantes une meilleure résistance aux stresses causés par les HP et les éléments-traces métalliques (Alarcón *et al.*, 2008; Liu & Dalpé, 2009; Dong *et al.*, 2014; Xun *et al.*, 2015; Mishra *et al.*, 2016). Cependant, les démarches expérimentales de l'ensemble de ces travaux sont limitées à la réalisation des combinaisons entre des espèces types de bactéries et de CMA, et leurs résultats sont restreins à l'observation des effets de ces combinaisons sur le rendement de dissipation des polluants dans le sol. À ce jour, on connaît peu de choses sur les interactions et la structure des communautés des CMA et les microorganismes qui leur sont associés dans les sites contaminés par les HP et les éléments traces métalliques.

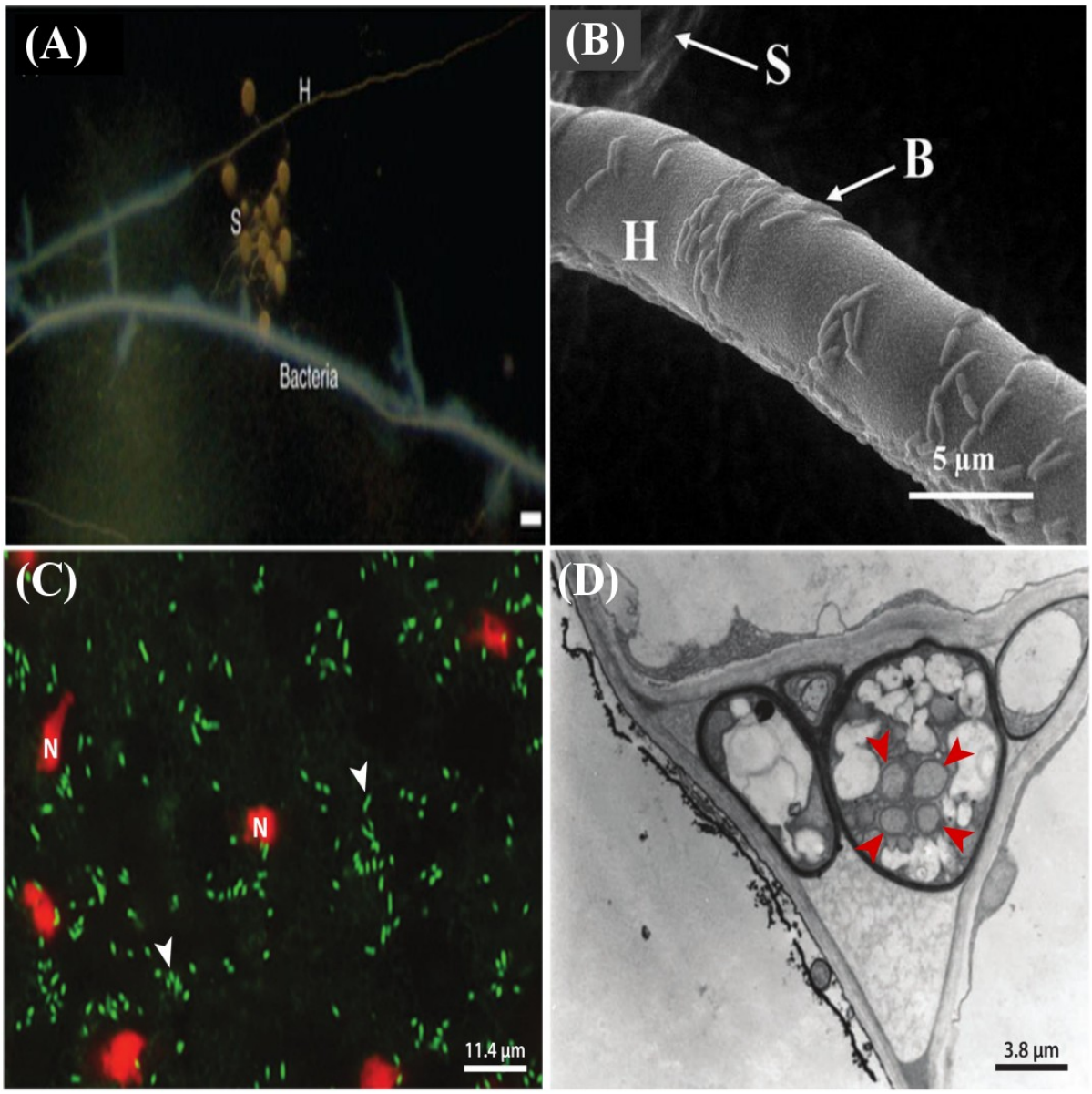


Figure 1. 2: Les structures microbiennes associées aux spores et mycéliums des CMA. (A) Observation microscopique des bio-filmes bactériens associés aux mycéliums des CMA (la barre d'échelle au-dessous de l'image représente 100 µm) (Lecomte et al., 2011). (B) Des structures bactériennes à la surface des hyphes des CMA observées par microscope électronique à balayage (Cruz & Ishii, 2012). (C) Image de microscopie confocale des structures bactériennes (les flèches blanches) à l'intérieur des spores de CMA (Bianciotto et al., 2000). (D) Image de microscopie électronique à transmission montrant des structures microbiennes (les flèches rouges) à l'intérieur des spores des CMA (Bonfante & Anca, 2009).

1.4. Description du site d'échantillonnage et mise en contexte de mon projet :

Les bassins de décantation des résidus pétroliers de l'ancienne raffinerie pétro-chimique de Varennes (Rive-Sud de Montréal, Canada) sont considérés parmi les sites les plus contaminés en hydrocarbures pétroliers dans la région de Montréal. D'après des analyses réalisées par notre laboratoire, les concentrations en hydrocarbures pétroliers dans ce site ont atteint des niveaux extrêmes. Par exemple, les concentrations de phénanthrène, anthracène et fluorène, des hydrocarbures aromatiques polycycliques (HAP) hautement cancérigènes, mesurées dans ce site dépassent les 4300 mg, 570 mg et 710 mg par 1 kg de sol, respectivement (voir Table S2.1 du chapitre 2). D'autre part, les normes admissibles par le MDDELCC pour ces trois composés dans les zones industrielles sont de l'ordre de 50 mg, 100 mg, 100 mg par 1 kg de sol [2].

Face aux coûts élevés et l'inefficacité des méthodes physiques et chimiques de décontamination, l'utilisation de méthodes biologiques pour la décontamination *in situ* de ces bassins semble le moyen le plus approprié pour la réhabilitation de ce site.

Mon projet de doctorat s'inscrit dans le cadre d'un projet d'envergure qui étudie la phytoremédiation du site pollué de Varennes en utilisant plusieurs cultivars du saule (*Salix* sp.) et des microorganismes connus pour leurs aptitudes à dégrader les résidus pétroliers. Cependant, lors de notre première visite à ce site, nous avons constaté que plusieurs espèces de plantes poussent spontanément dans les basins de décantation de cette ancienne raffinerie pétro-chimique. Selon Desjardins *et al.* (2014), il y a 23 espèces de plantes qui poussent

[2]Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques (MDDELCC); http://www.mddelcc.gouv.qc.ca/

spontanément dans ce site pollué de Varennes. Par conséquent, ces observations ont soulevé les questions suivantes :

1- Comment ces espèces de plantes, dont certaines ne sont pas connues dans le domaine de phytoremédiation, arrivent-elles à s'installer dans un site aussi contaminé par les hydrocarbures pétroliers toxiques?
2- Quels sont les mécanismes qui permettent à ces plantes de tolérer ou dégrader les résidus pétroliers?
3- Ces plantes bénéficient-elles de la flore microbienne indigène de ce site? Si l'établissement de ces plantes est soutenu par la flore microbienne de ce site, quelles sont les communautés microbiennes que l'on peut trouver en association avec ces plantes?
4- Les CMA, font-ils partie des communautés microbiennes associées à ces plantes ?
5- Étant donné que les CMA recrutent aussi des microorganismes à la surface et à l'intérieur de leurs spores et mycélia, quelle est la structure des communautés microbiennes associée aux CMA dans un site contaminé par les HP? La structure des communautés microbiennes associées aux CMA varie-t-elle en fonction des niveaux de contaminations et espèces de plantes? La structure des communautés microbiennes associée aux CMA est-elle différente de celle de la rhizosphère et les racines de plantes?

Dans ce contexte, mon projet de doctorat a été mis en place pour tenter de répondre à certaines de ces nombreuses questions. Mes objectifs dans ce projet sont :

i. d’étudier la diversité et les structures des communautés des CMA et les microorganismes qui leur sont associés dans des sols contaminés par les HP;

ii. d’étudier la variation des structures des communautés des CMA ainsi que celle des microorganismes qui leur sont associés par rapport au niveau de concentration en HP et à aux espèces de plantes hôtes;

iii. d’étudier l'impact des CMA sur la structure des communautés microbiennes associées à leurs spores dans ce site contaminé;

iv. de comparer les communautés microbiennes des racines et des sols contaminés par les HP avec celles associées avec les CMA.

1.5. Résumé de la démarche expérimentale

Afin de réaliser mon projet de doctorat, j'ai choisi trois espèces de plantes, *Solidago canadensis*, *Populus balsamifera* et *Lycopus europaeus* qui poussent spontanément dans trois bassins de décantation ayant des niveaux de contamination en HP différents (voir Table S3.10. du chapitre 3). Mon choix s'est porté sur ces trois espèces de plantes en raison de leur abondance, leur présence dans les trois bassins et leurs capacités à former des associations symbiotiques avec les CMA. Pour chacune des espèces de plantes et dans chaque bassin, j'ai échantillonné trois individus (individus entiers avec le sol entourant les racines) par espèce et par bassin en totalisant ainsi 27 échantillons (3 espèces de plantes × 3 bassins × 3 individus). Une fois au laboratoire, une partie des échantillons du sol a été utilisée dans l'extraction des spores de CMA, tandis qu'une partie des racines de plantes a été utilisée dans les observations microscopiques et l'extraction des propagules intra-racinaires. Les sous-échantillons restants des sols et des racines ainsi que les spores et les propagules extraites ont été soumis à des

techniques de microscopie, d'amplification de l'ADN génomique total (*Whole Genome Amplification*), de PCR, de clonage et de séquençage de Sanger ou séquençage à haut débit. Ensuite, des analyses bio-informatiques et statistiques ont été réalisées afin d'évaluer les effets des paramètres biotiques et abiotiques sur les communautés des CMA et les microorganismes qui leur sont associés.

Mise en contexte du chapitre 2 - Quel est la biodiversité des bactéries associées aux CMA dans un site hautement contaminé par les hydrocarbures pétroliers ?

Il est admis que les CMA établissent des associations symbiotiques ou parasitiques avec la flore microbienne présente dans la rhizosphère des sols agricoles. Cependant, on connaît peu de choses sur la biodiversité des bactéries associées aux CMA dans les sites contaminés par les HP et encore moins dans les propagules intra-racinaires. À cet effet, dans ce chapitre, nous avons développé une approche originale basée sur la microdissection des racines mycorhizées pour étudier la diversité bactérienne associée aux propagules intra-racinaires des CMA extraites à partir des racines de la plante *Solidago rugosa* échantillonnées dans le site pollué de Varennes.

Les résultats de ce chapitre ont fait l'objet d'une publication dans la revue *FEMS Microbiology letter* en 2014.

Ma contribution dans cet article est la participation à la mise en place de l'expérience et la réalisation de la totalité des travaux d'échantillonnage et de laboratoire. J'ai aussi participé à la rédaction et j'ai fait toutes les illustrations présentées dans l'article.

Chapitre 2 - Bacteria associated with arbuscular mycorrhizal fungi within roots of plants growing in a soil highly contaminated with aliphatic and aromatic petroleum hydrocarbons

Bachir Iffis, Marc St-Arnaud and Mohamed Hijri

Institut de Recherche en Biologie Végétale, Département de Sciences Biologiques, Université de Montréal, 4101 Sherbrooke est, Montréal (Québec) H1X 2B2, Canada

Published in : FEMS Microbiology Letters, August 4, 2014. 358: 44–54. doi: 10.1111/1574-6968.12533

Author contributions : Conceived and designed the experiments: BI MH. Performed the experiments: BI. Analyzed the data: BI. Contributed reagents/materials/analysis tools: BI. MH. Wrote the paper: BI MH MSA.

2.1. Abstract

Arbuscular mycorrhizal fungi (AMF) belong to phylum Glomeromycota, an early divergent fungal lineage forming symbiosis with plant roots. Many reports have documented that bacteria are intimately associated with AMF mycelia in the soil. However, the role of these bacteria remains unclear and their diversity within intraradical AMF structures (AMF structures inside roots) has yet to be explored. We aim to assess the bacterial communities associated within intraradical propagules (vesicles and intraradical spores) harvested from roots of plant growing in the sediments of an extremely petroleum hydrocarbon-polluted basin. *Solidago rugosa* roots were sampled, surface-sterilized, and microdissected. Eleven propagules were randomly collected and individually subjected to whole-genome amplification, followed by PCRs, cloning, and sequencing targeting fungal and bacterial rDNA. Ribotyping of the 11 propagules showed that at least five different AMF OTUs could be present in *S. rugosa* roots, while 16S rRNA ribotyping of six of the 11 different propagules showed a surprisingly high bacterial richness associated with the AMF within plant roots. Most dominant bacterial OTUs belonged to *Sphingomonas* sp., *Pseudomonas* sp., *Massilia* sp., and *Methylobacterium* sp. This study provides the first evidence of the bacterial diversity associated with AMF propagules within the roots of plants growing in extremely petroleum hydrocarbon-polluted conditions.

2.2. Keywords

Arbuscular mycorrhizal fungi; bacteria; petroleum hydrocarbons; *Solidago rugosa*; 16S and 18S rRNA genes.

2.3. Introduction

Arbuscular mycorrhizal fungi (AMF) belong to an ancient group of plant-inhabiting fungi that form symbiotic associations. The arbuscular mycorrhizal association is among the oldest symbioses between plants and fungi on earth, and it has been dated back to Ordovician (*c.* 450 million years) using fossil records and molecular clocks (Simon *et al.*, 1993; Redecker *et al.*, 2000). AMF belong to phylum Glomeromycota, an early diverging fungal lineage of ecologically and economically relevant microorganisms. Glomeromycota promote plant growth by enhancing mineral uptake, in particular phosphorus, and protect plants against soil-born pathogens (St-Arnaud & Elsen, 2005; Smith & Read, 2008; Ismail *et al.*, 2011; Ismail *et al.*, 2013).

AMF are widely distributed and can be found in all ecosystems on earth where plants are able to grow. Therefore, they are usually considered to be generalist plant symbionts, as their diversity is limited to between 200 and 300 described species (Öpik *et al.*, 2010; Redecker *et al.*, 2013). The analysis of a large dataset of 14 961 AMF nucleotide sequences retrieved from 111 studies showed that geographic distance, soil temperature and moisture, and plant community type were each significantly related to AMF community structure, but these factors explain only a small amount of the observed variance in the meta-analysis data (Kivlin *et al.*, 2011). Soil pollutants have been considered as potentially major factors affecting AMF diversity (Vallino *et al.*, 2006; Bedini *et al.*, 2010; Long *et al.*, 2010; Hassan *et al.*, 2011). The role of AMF in polluted soils has been widely studied, and several studies revealed that AMF promote phytoremediation and enhance plant tolerance against trace metals and petroleum hydrocarbon pollutants (Bedini *et al.*, 2010; Aranda *et al.*, 2013; Hassan *et al.*, 2013). For example, Bedini *et al.* (2010) found nine AMF OTUs associated with plants

spontaneously growing in trace metal-polluted ash disposal site containing extreme concentrations of Cu, Pb, and Zn.

In all types of soils where AMF are found, their growth is not limited to plant roots; their mycelia extend beyond roots, exploring a larger volume of soil, and producing extraradical hyphae and asexual multinucleate spores. The mycelia of AMF are typically coenocytes that lack septa, allowing cytoplasm, nuclei, and organelles to move freely within hypha (Marleau *et al.*, 2011).

In natural ecosystems, numerous bacterial taxa are closely associated with AMF mycelia where they colonize the surface of extraradical hyphae and spores on which they can form biofilm-like structures (Scheublin *et al.*, 2010; Lecomte *et al.*, 2011; Cruz & Ishii, 2012). Both culture-dependent and culture-independent methods have observed and identified several bacterial taxa belonging to *α*-, *β*-, and *γ*-*Proteobacteria* and *Firmicutes* from the surface of mycelia of many AMF species (Roesti *et al.*, 2005; Bonfante & Anca, 2009; Scheublin *et al.*, 2010; Lecomte *et al.*, 2011). In some AMF taxa, bacteria were also shown to live in the cytoplasm as endobacteria [reviewed in (Bonfante & Anca, 2009)]. Using microscopy, Bianciotto *et al.* (1996) found that an individual spore of the AMF *Gigaspora margarita* can harbor up to 250 000 bacterial cells in its cytoplasm. However, Jargeat *et al.* (2004) attempted to cultivate the endobacterium *Candidatus Glomeribacter gigasporarum* living inside *G. margarita* using 19 different culture media without notable success. This supports the idea that some endobacteria could be obligate biotrophs that are not able to grow without AMF, which themselves require a host plant to complete their life cycle (Jargeat *et al.*, 2004). Interaction of AMF and bacteria brings another level of complexity to diversity and function of the mycorrhizal symbiosis. Thus, some authors hypothesize that plants, AMF, and bacteria can be

considered as tripartite associations resulting in a consortium that promotes plant growth (Bonfante & Anca, 2009). However, the potential roles and infection mechanisms of these bacteria, in particular the endobacteria, are still poorly understood. In addition, the diversity of these associated bacteria has not been explored in polluted soils, neither in extraradical spores nor inside the mycorrhizal roots.

The objective of this study was therefore to describe the bacterial diversity associated with AMF propagules (vesicles and intraradical spores) extracted from the roots of a plant species growing spontaneously in a decantation basin extremely polluted with petroleum hydrocarbons. To do so, we microdissected mycorrhizal *Solidago rugosa* roots harvested from a polluted site to isolate intraradical AMF propagules. Then, each propagule was subjected to whole-genome amplification (WGA). *Bacterial* diversity was assessed using cloning and sequencing of the 16S rRNA genes. This approach allowed us to profile the bacteria closely associated with the AMF while reducing the additional diversity of soil bacteria, which can be randomly attached to the surface of AMF extraradical mycelia. 18S rRNA ribotyping was also performed on WGA products to assess AMF taxonomic diversity.

2.4. Materials and methods

2.4.1. Site of study, harvesting, and preparation of samples

Sampling occurred on the site of a former petrochemical plant located on the south shore of the St-Lawrence River near Montreal, Quebec, Canada (45°41′55.3″ N 73°25′45.0″ W). Various plants species were spontaneously growing in an open-air sedimentation basin in which petroleum hydrocarbon wastes were dumped for many decades.

Polycyclic aromatic hydrocarbon (PAH) and alkane (C10–C50) concentrations in the basin are shown in Supporting Information, Table S2.1. PAH and alkane concentrations exceeded by hundreds of time the standards for reuse defined by the government of Quebec for industrial areas. Three individual plants of *S. rugosa* Mill. growing in the basin sediments were collected in October 2011. Plant roots were cut, were washed several times with tap water to remove rhizospheric soil, and were cut again into pieces *c.* 1 cm long. A subsample of the roots from each plant was stained for microscopy observations, as described below. Under sterilized laminar flow hood, the remaining root pieces were washed in sterilized water containing a few drops of Tween 80 to favor removal of the petroleum hydrocarbons attached to the root surface and then in each of the following surface disinfecting treatments: ethanol 90 % for 10 s, commercial sodium hypochlorite 5 % for 2 min, chlromine-T 4% for 10 min, and streptomycin 0.03 % for 30 min. After the last treatment, the roots were washed several times in sterile distilled water and stored in 1.5 mL microtubes prior to microdissection. To test the efficiency of the surface sterilization procedure, *c.* 20 root fragments were transferred to Petri dishes containing TSA or PDA media and incubated for 2 weeks to check for the presence of bacteria or fungi able to grow in these media.

2.4.2. Estimation of mycorrhizal root colonization

The roots of *S. rugosa* were cleared in a solution of KOH 10 % at 80 °C for 1 h, washed several times in deionized water, and stained in a 1 % acid fuchsin solution at 60 °C for 1 h. The roots then were washed, cut to small fragments, and mounted on microscope slides using glycerol 60 % as a mounting medium. The percentage of AMF root colonization was determined under the microscope using the grid-line intersect method (Giovannetti & Mosse, 1980).

2.4.3. Root microdissection, extraction of AMF propagules, and scanning electron microscopy (SEM)

Disinfected root pieces were soaked in a filtered (0.22 µm) mixture of enzymatic solution of 2 % pectolyase (Sigma-Aldrich) and 3 % cellulase (Sigma‑Aldrich) (w/v) in sterile water at 30 °C for 1 h to digest root cell walls. Root pieces were rinsed and transferred to Petri dishes in which they were microdissected under a stereomicroscope positioned into a horizontal clean bench, using thin sterile forceps and needles. Fourteen fungal propagules were randomly collected and individually put in 0.2 mL microtubes containing 2 µL of sterile water and kept at –20 °C until use. Because clear discrimination between vesicles and intraradical spores that some AMF are able to produce requires destructive examination of the cell wall at high magnification, we used the term 'AMF intraradical propagules' to designate both structure types. We also collected an uncolonized root tip sample that was used as a control for assessing bacterial endophytes colonization within *S. rugosa* roots. Workflow of the experimental approach is summarized in Figure 2.1. To visually confirm the presence of bacteria on the surface of AMF propagules, root pieces were prepared for SEM following the protocol described in Bozzola and Russell (1992). A Quanta 200 3D (FEI, Burlington, MA) SEM was used to visualize samples and acquire images.

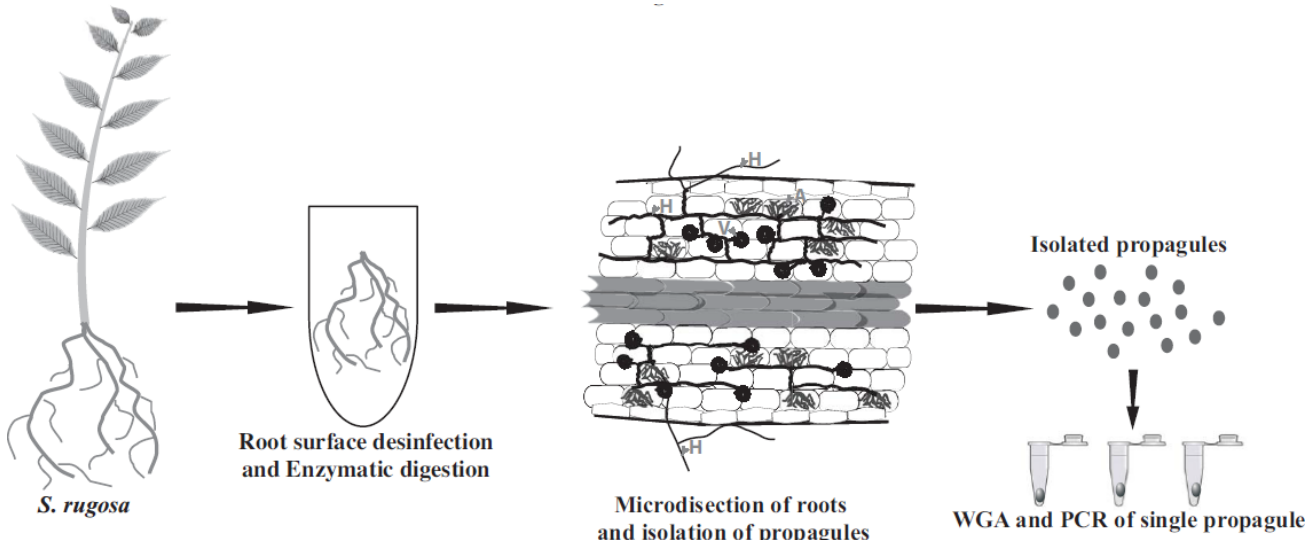


Figure 2.1. Workflow of the experimental approach used in this study, consisting of the collection of root samples from the field, root sterilization, digestion of cell wall, and microdissection.
Isolated propagules were subjected to WGA, followed by PCRs targeting the fungal 18S and bacterial 16S rRNA genes, cloning, and sequencing.

2.4.4. WGA and PCR

WGA reactions were performed directly on each individual propagule and on uncolonized root tip using the Illustra GenomiPhi HY DNA Amplification Kit (GE Healthcare Life Sciences, QC, Canada) according to the manufacturer's instructions. WGA reactions were carried out in order to have sufficient quantity of DNA for doing several PCR amplifications. All WGA products were stored at –20 °C until use. PCRs were then performed using WGA products as DNA template to amplify 18S rRNA gene fragments using AML1 and AML2 primers (Lee *et al.*, 2008), to identify AMF sequences. Nested PCRs were also performed with the primer pair AM1 and NS31 (Simon *et al.*, 1992; Lee *et al.*, 2008) on the propagules for which no amplification occurs using the AML1/AML2 primers. PCR amplifications were also

performed using the 16S rRNA primer pair 27f and 1495r (Bianciotto *et al.*, 1996) to assess bacterial diversity associated with the AMF propagules.

PCRs contained PCR buffer, 0.5 µM of each primer, 0.2 mM of dNTP, 1 µL of WGA product, and 1 U of *Taq* polymerase (QIAGEN, Toronto, ON, Canada) in a total volume of 40 µL. PCRs were run on a MasterCycler ProS thermocycler (Eppendorf, Mississauga, ON, Canada) under the following program: an initial denaturation at 95 °C for 5 min followed by 38 cycles of 95 °C for 30 s, 54 °C for 30 s, and 72 °C for 90 s, and a final elongation at 72 °C for 10 min. PCR products were separated by agarose gel (1 %) electrophoresis, stained with GelRed, and visualized under UV light using a Gel-Doc system (Bio-Rad Laboratories, Mississauga, ON).

2.4.5. Cloning and sequencing

Cloning reactions were performed individually on 16S and 18S rDNA PCR products. The ligation reactions were performed in a volume of 10 µL using pGEM-T Easy Vector System II kit containing chemically competent JM109 *Escherichia coli* cells (Promega, ThermoFisher, Ottawa, ON, Canada) according to recommendations of the manufacturer. Bacterial colonies were screened by PCR with T7 and SP6 universal primers (Hassan *et al.*, 2011). Positive clones that contained inserts were sent for sequencing at the McGill University and Genome Quebec Innovation Centre (Montreal, QC). Sequences were deposited in GenBank under accession numbers KJ809141–KJ809555.

2.4.6. Data analyses

Clustering of bacterial sequences was performed in GENEIOUS version 6 (Biomatterts Limited, Auckland, New Zealand), and OTUs were defined at 98 % similarity. Singleton

sequences were kept and used in a different analysis. Rarefaction analyses were performed in R version 3.0.1 software using the VEGAN package (http://www.r-project.org). The estimator of sample coverage was calculated using INEXT online (http://chao.stat.nthu.edu.tw/blog/software-download/); (Chao & Jost, 2012).

For phylogenetic analysis, sequence similarities of AMF 18S rRNA genes were obtained from MaarjAM (Öpik *et al.*, 2010) and GenBank databases. Choanoflagellate species *Monosiga brevicollis* and *M. ovata* were used as an out-group for this analysis. A phylogenetic tree was generated using a neighbor-joining approach with 1000 bootstrap resamplings using the MEGA version 5.10 software (Tamura *et al.*, 2011).

2.5. Results and discussion

2.5.1. Diversity of AMF in plant roots

Microscopic examination of *S. rugosa* roots, a plant species spontaneously growing in the sediments of a decantation basin containing very high concentrations of aliphatic and aromatic petroleum hydrocarbons (Table S2.1), showed typical AMF vesicles and hyphae (Figure S2.1). The roots showed mycorrhizal colonization with frequency of 70 %. As it is not possible to distinguish AMF species based on microscopic examination of roots, we used WGA, PCR, and cloning of the 18S rRNA gene to assess AMF diversity within roots of *S. rugosa*. Among the 14 propagules analyzed, 11 led to successful PCR amplification products that were subsequently cloned and provided a total of 41 clones. Clone sequences, which were amplified using AML1/AML2 primers, generated a sequence length of *c.* 800 bp, except those from propagule 5 which were amplified using AM1/NS31 primers and were *c.* 550 bp. The

number of clones, sequence lengths, and BLASTN similarity results for each propagule are given in Table 2.1.

Table 2.1. Arbuscular mycorrhizal fungal taxa found in propagules isolated from Solidago rugosa roots, based on 18S rRNA gene sequencing

Propagules	*Number of clones*	*Fragment length (bp)*	*Most closely related taxa*	*Accession number*	*% of identity*
1	3	800-801	*Diversispora eburnea VTX00060*	AM713429	99
2	1	798	*Archaeospora schenckii VTX00245*	FR773150	96
3	6	795-796	*Glomus sp. VTX00419*	GQ140619	95-97
4	1	805	*Claroideoglomus Torrecillas12b Glo G1 VTX00193*	HE614988	99
5	4	539-560	*Claroideoglomus Torrecillas12b Glo G4 VTX00056*	HE615005	98
6	2	798	*Glomus irregulare VTX00114*	FN600536	99
7	4	805-806	*Claroideoglomus Early-1 VTX00056*	JN252440	96-97
8	9[1]	794-805	*Claroideoglomus Early-1 VTX00056*	JN252440	97
			Spizellomyces palustris	FJ827665.1	95
9	3	805-806	*Claroideoglomus Torrecillas12b Glo G1 VTX00193*	HE614988	99
10	6	799-804	*Diversispora eburnea VTX00060*	AM713429	98-99
11	2	789-800	*Diversispora eburnea VTX00060*	AM713429	99

[1] Three clones matched with *Claroideoglomus* Early-1 VTX00056, while, six matched with *Spizellomyces palustris* which is a Chitridiomycete.

Phylogenetic analysis showed a high AMF species diversity colonizing *S. rugosa* roots (Figure 2.2). The 18S rRNA gene sequences from all propagules clustered within taxa belonging to four different AMF families: *Claroideoglomeraceae* (five propagules), *Diversisporaceae* (three propagules), *Glomeraceae* (one propagule), and *Archaeosporaceae* (one propagule). Clones from propagule 3 showed 95–97 % identity with an unidentified AMF sequence (accession number: GQ140619.1) closely related to *Glomeraceae* which was found in trace metal-contaminated soil in China (Long *et al.*, 2010). Interestingly, clones from

propagule 6 showed 98–99 % similarity with AMF sequences VeGlo8 and VeGlo10, closely related to *Rhizophagus irregularis* (previously named as *Glomus intraradices*), which were found in roots of plants growing in an extreme site polluted with Cu, Pb, and Zn in Venice, Italy (Bedini *et al.*, 2010).

No sequence matching with *Paraglomeraceae*, *Gigasporaceae*, or *Acaulosporaceae* was found in this study. However, this was expected as most AMF taxa belonging to these families are not known to form vesicles and spores within plant roots (Oehl *et al.*, 2011). Arbuscular mycorrhizal fungal diversity in trace metal polluted soil has been extensively studied. For example, Vallino *et al.* (2006) identified 13 taxonomic units belonging to *Glomeraceae*, *Diversisporaceae*, and *Gigasporaceae* from roots of *S. gigantea* naturally growing in a trace metal-contaminated site in northern Italy. Long *et al.* (2010) studied AMF diversity from roots and rhizospheric soil of five plants species growing in trace metal-contaminated soil and also identified species belonging to *Glomeraceae*, *Claroideoglomeraceae*, *Acaulosporaceae*, and *Archaeosporaceae*. In another study, Hassan *et al.* (2011) showed that the community structure of AMF associated with *Plantago major* plants was determined by trace metal concentrations in the soil and dominated by *Funneliformis mosseae* in polluted sites. In contrast, although several studies showed that AMF may directly or indirectly influence phytoremediation of petroleum hydrocarbons (Wu *et al.*, 2009; Gao *et al.*, 2010; Gao *et al.*, 2011a; Gao *et al.*, 2011b; Aranda *et al.*, 2013), the effect of these compounds on AMF diversity remains unclear.

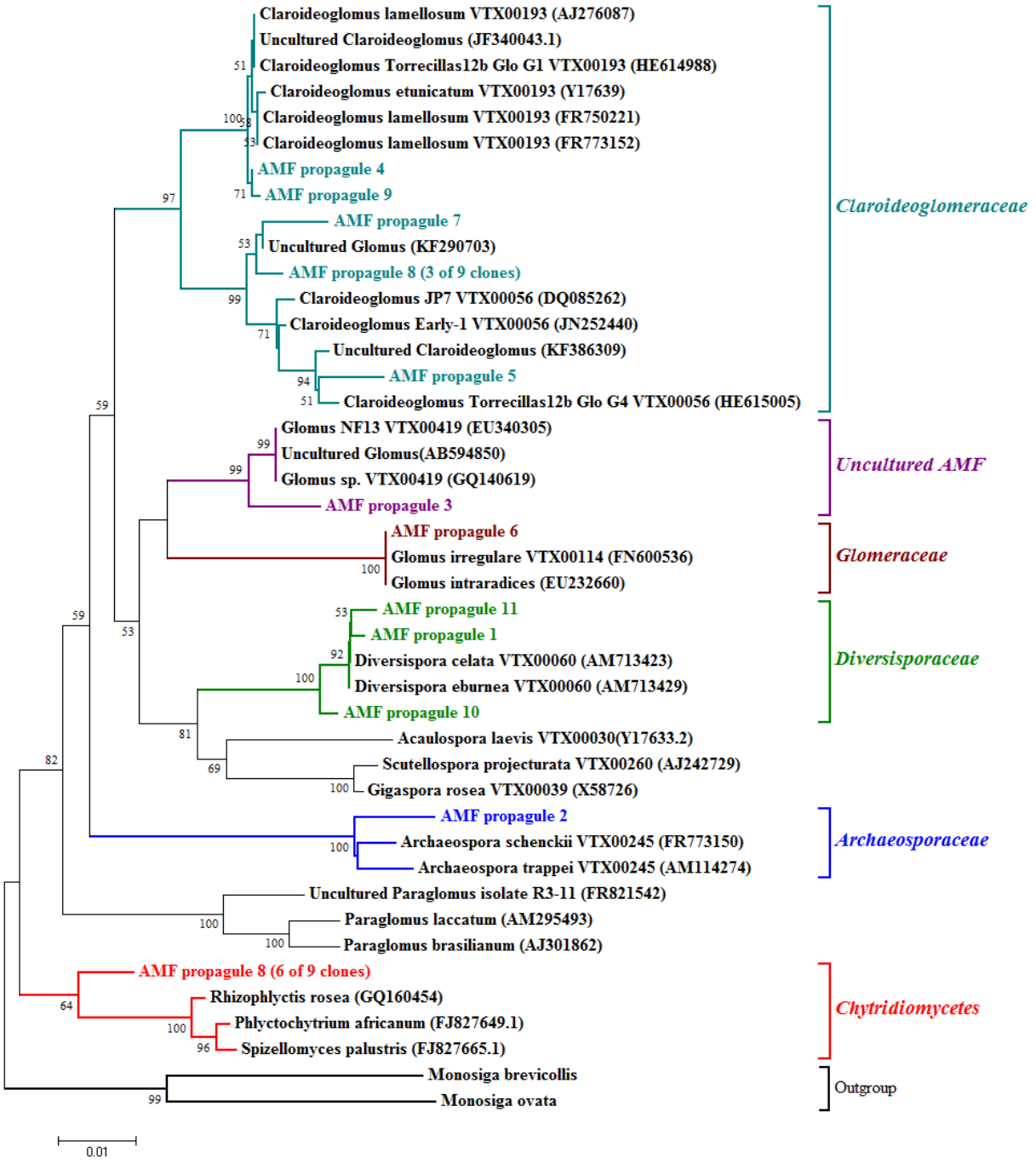


Figure 2.2. Neighbor-joining tree of the 18S rRNA gene of the consensus sequences of clones obtained from each of the 11 propagules analyzed in our study except for propagule 8 in which two consensus sequences of different taxonomic origin were found. The tree shows the different AMF families in which each propagule consensus sequence clustered, and it also shows that six clones of propagule 8 clustered within Chytridiomycetes. Bootstrap values lower than 50 % of 1000 replicates are not shown on the branches.

Interestingly, six of nine 18S rRNA gene clones retrieved from propagule 8 matched with Chytridiomycota species, showing 95 % identity to *Spizellomyces palustris* and *Rhizophlyctis rosea* (accession numbers: FJ827665 and GQ160454, respectively). This is not surprising because Chytridiomycota are commonly found in wet soils and can be endophytic or pathogens of plant roots (Jobard *et al.*, 2010). However, little is known about the interaction between AMF and Chytridiomycota. Only a few studies from the 1970s and 1980s hypothesized that some chytrids such as *Rhizidiomycopsis* sp., *Phlyctochytrium* sp., and *Spizellomyces* sp. could be hyperparasites of AMF and may negatively affect the mycorrhizal symbiosis (Ross & Ruttencutter, 1977; Schenck & Nicolson, 1977; Sylvia & Schenck, 1983). However, Paulitz and Menge (1984) also proposed that Chytridiomycota could be saprotrophs of decaying AMF structures and reported that *Spizellomyces punctatum* infected mainly nongerminated or dead spores of *Gigaspora margarita*. Tzean *et al.* (1983) suggested that *Phlyctochytrium kniepii* may be vector for transmitting bacteria such as *Spiroplasma-like organisms* to the cell wall and cytoplasm of AMF spores. It has also been reported that some AMF spores could be infected by other fungi belonging to *Ascomycota* (Hijri *et al.*, 2002).

2.5.2. Bacterial diversity associated with AMF propagules

Among the 11 AMF propagules successfully identified, six propagules belonging to different AMF species (propagules 1‑6) were used to assess the associated bacterial diversity. A total of 428 clones were sequenced from the six propagules and from a control noncolonized root tip. Of these clones, 53 sequences matched with plant chloroplastic genes and were removed from the analysis. The remaining 375 sequences matched with bacterial 16S rRNA genes and were clustered at 98 % sequence identity resulting in 27 OTUs and 23 additional

singleton sequences. The bacterial genera obtained from the different OTUs groups and singletons are shown in Table 2.2. Rarefaction analyses were performed on each propagule sequence dataset (Figure 2.3). The highest coverage was obtained from the propagules 2, 5, and 6 with recovery of 91.7 %, 92.6 %, and 90.4 %, respectively, while the lowest coverage was obtained from the propagule 4 with 66.2 %. Propagules 1 and 3 showed intermediate values with 80.4 % and 87 %, respectively. The control root tip showed a saturated rarefaction curve with 100 % of sample coverage and was represented only by two OTUs (Figure 2.3). The most dominant bacterial OTUs associated with AMF propagules belonged to *Sphingomonas* sp. (28.2 %), *Pseudomonas* sp. (15.7 %), *Massilia* sp. (14.4 %), *Methylobacterium* sp. (11.7 %), and unidentified bacterium (9.8 %). Other OTUs belonging to *Bradyrhizobium* sp., *Bacillus* sp., *Bosea* sp., and *Paenibacillus* sp. were found at lower frequencies (Figure S2.2). The highly abundant bacterial OTUs were observed in almost all propagules with variable frequencies between the propagules. Those that were less abundant were only found in specific propagules. This may be due to the sampling effort, which did not cover all the bacterial diversity associated with propagules, or perhaps there is a specific affinity of bacteria for AMF species. For example, *Pseudomonas* sp. was the dominant OTU in propagules 1 and 5 with proportions of 33 % and 42 %, respectively, but it was not detected in propagule 2 and was only found at low frequencies in other propagules. The same is true for *Sphingomonas* sp. which was detected as the dominant OTU for propagules 2, 4, 5, and 6, while only one clone belonging to this taxon was found in propagule 1 (Figure S2.3). Interestingly, the bacterial OTU richness found in the root sample was extremely low in comparison with that found in AMF propagules as it was represented by only two OTUs, which were also found in AMF propagules, and these belonged to *Pseudomonas* sp. (14

clones) and *Streptococcus* sp. (15 clones). *Pseudomonas* sp. was largely dominant (45 clones from 5 of the 6 AMF propagules), while *Streptococcus* sp. was found in propagules 1 and 2 only (with seven clones in total). *Pseudomonas* spp. were reported as an endophytic as well as AMF-associated bacterium, able to promote both plant growth and the symbiotic association between AMF and plants (Strobel & Daisy, 2003; St-Arnaud & Elsen, 2005; Scheublin *et al.*, 2010; Gaiero *et al.*, 2013). The low bacterial richness found in the root may be due to the effect of the surface disinfection of root fragments or to the fact that meristematic cells forming a significant part of the root tip are mainly exempt of bacterial colonization.

Bacteria found to be associated with AMF propagules and within roots may either colonize the AMF cytoplasm or may be attached to their external surface, although both situations were reported to occur for some bacteria (Bonfante & Anca, 2009). To test whether bacteria were attached to the surface of the intraradical AMF propagules, microdissected mycorrhizal roots previously disinfected, and processed under sterile conditions were examined using SEM. Coccoid bacteria and biofilm-like structures attaching to the surface of AMF propagules were clearly visible inside the cortex of *S. rugosa* roots (Figure 2.4). The size of these bacterial cells ranged from 0.5 to 1.5 µm.

Table 2.2. Detection frequencies and identity of bacterial genera associated with AMF propagules isolated form *Solidago rugosa* based on 16S rRNA gene sequencing OTUs were clustered at 98 % of sequence identity; then, OTUs showing similarities with the same bacterial genera were grouped together. P means propagule; R means roots.

Bacterial taxa	P-1	P-2	P-3	P-4	P-5	P-6	R	Fragment length	Percentage of identity	Accession Numbers
Sphingomonas sp.	1	39	26	14	11	15	-	664-1463	91%-98.5%	EF061133, JQ659520, JX566637, AY749436, HF558376, AF181571, AB109749, JX879745
Pseudomonas sp.	22	-	1	1	11	10	14	325-1538	98%-99%	AM421016, FJ889609, KF011692
Massilia sp.	5	26	16	5	1	1	-	827-1505	94%-95%	JX566602
Methylobacterium sp.	4	3	32	-	-	5	-	604-1487	92%-99%	AB698713, FJ025133, CP001029, JF905617
Uncultured bacterium	11	4	17	-	-	5	-	791-1506	96%-99%	HE576045, AY672523, JF429334, DQ129631, JX647723, HE798198, JN023771, JX271950, FJ984639, HM845051, JQ769980, GU563747
Streptococcus sp.	2	5	-	-	-	-	15	876-1517	92%-99%	FR875178, FQ312041
Bosea sp.	7	4	1	1	-	-	-	846-1453	97%	DQ440827
Afipia sp.	1	1	-	1	-	4	-	842-1453	99%	DQ123622
Brevundimonas sp.	-	-	2	2	-	1	-	912-1430	98%	KC494321
Bradyrhizobium sp.	4	-	-	-	-	-	-	1461-1486	98%	FJ390912
Paenibacillus sp.	3	-	-	-	-	-	-	1034-1517	97%	JX566644
Lysobacter sp.	-	1	2	-	-	-	-	946-1002	97%-99%	JQ746036
Acinetobacter sp.	-	-	-	2	1	-	-	868-1504	99%	HE651920 AB099655
Stenotrophomonas sp.	-	-	-	-	2	-	-	890-1514	99%	KF150351
Propionibacterium sp.	1	-	-	1	-	-	-	1495-1499	99%	CP003877
Agrobacterium sp.	2	-	-	-	-	-	-	816-898	99%	AY174112
Legionella sp.	1	-	-	1	-	-	-	883-938	95%-97%	AM747393 JF779686
Pseudacidovorax sp.	-	-	2	-	-	-	-	951-1000	97%-98%	HQ834240
Azospirillum sp.	-	-	-	1	-	-	-	869	95%	AP010946
Bacillus sp.	-	1	-	-	-	-	-	1521	97,50%	JQ435679
Lactobacillus sp.	-	-	1	-	-	-	-	1540	99%	EU855223
Leptothrix sp.	1	-	-	-	-	-	-	1238	97,80%	AF385534
Pseudoxanthomonas sp.	1	-	-	-	-	-	-	929	98,70%	DQ337597
Total	66	84	100	29	26	41	29			

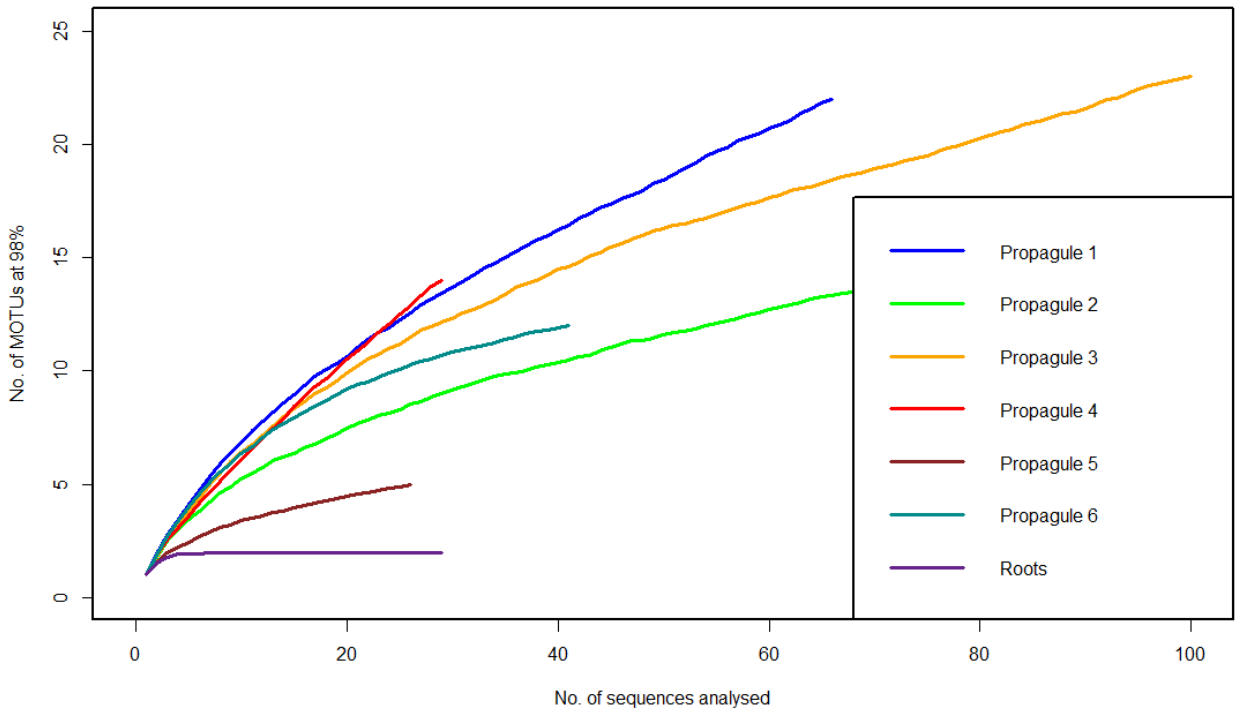


Figure 2.3. Rarefaction curves of the bacterial OTUs associated with AMF propagules and roots. OTUs were assigned at 98 % of sequence similarity. The estimation of sample coverage for propagules 1–6 and the uncolonized roots were respectively 80.4 %, 91.7 %, 87 %, 66.2 %, 92.2 %, 90.4 %, and 100 %.

Based on 16S rRNA genes, most bacterial OTUs retrieved from AMF propagules matched with bacterial taxa already reported to attach on the surface of spores and extraradical mycelia of AMF in soil (Roesti *et al.*, 2005; Scheublin *et al.*, 2010; Lecomte *et al.*, 2011). Our data support that these bacteria are not only able to interact with the external AMF mycelia in soil, but they also colonize the AMF propagules inside roots. However, the effect of these bacteria on AMF remains unclear. Dominant bacterial taxa found in this study, belonging to genera *Sphingomonas*, *Pseudomonas*, *Massilia*, and *Methylobacterium*, were previously identified in hydrocarbon-polluted soils and were shown to be involved in biodegradation of PAHs (Dennis & Zylstra, 2004; Van Aken *et al.*, 2004; Zhou *et al.*, 2006; Ní Chadhain & Zylstra, 2010; Zhang *et al.*, 2010). Interestingly, the nondominant taxa found here, which

belong to genera *Bosea*, *Brevundimonas*, *Bradyrhizobium*, and *Paenibacillus*, have been reported elsewhere to improve mycorrhizal colonization of roots and plant nutrient uptake (Frey-Klett *et al.*, 2007; Tarkka & Frey-Klett, 2008; Bonfante & Anca, 2009). The nondominant taxa we have found have also been considered to be members of a group called *mycorrhiza helper bacteria* (MHB), which includes phosphate-solubilizing bacteria (PSB) and nitrogen-fixing bacteria (Garbaye, 1994; Marschner & Timonen, 2006). The relationship between MHB and AMF has not yet been investigated in detail, but it has been suggested that MHB obtain their carbon resources from AMF hyphae (Bonfante & Anca, 2009) and that in return, these bacteria produce signaling compounds that can enhance the AMF stimulation of root exudates (Barea *et al.*, 2005). It has also been reported that *Paenibacillus* sp. can support the growth and sporulation of the AMF *Rhizophagus irregularis* (formerly named *G. intraradices*), independently from the presence of the plant (Hildebrandt *et al.*, 2006), although this report was controversial. Our results also support that intraradical AMF propagules can also harbor obligate endobacteria living inside AMF spores. For example, the clone *G15GN* sequenced from the propagule 1 (accession number KJ809239) showed a sequence identity of 97 % with an obligate endobacterium (accession number FJ984641) found in the cytoplasm of an AMF species (Naumann *et al.*, 2010).

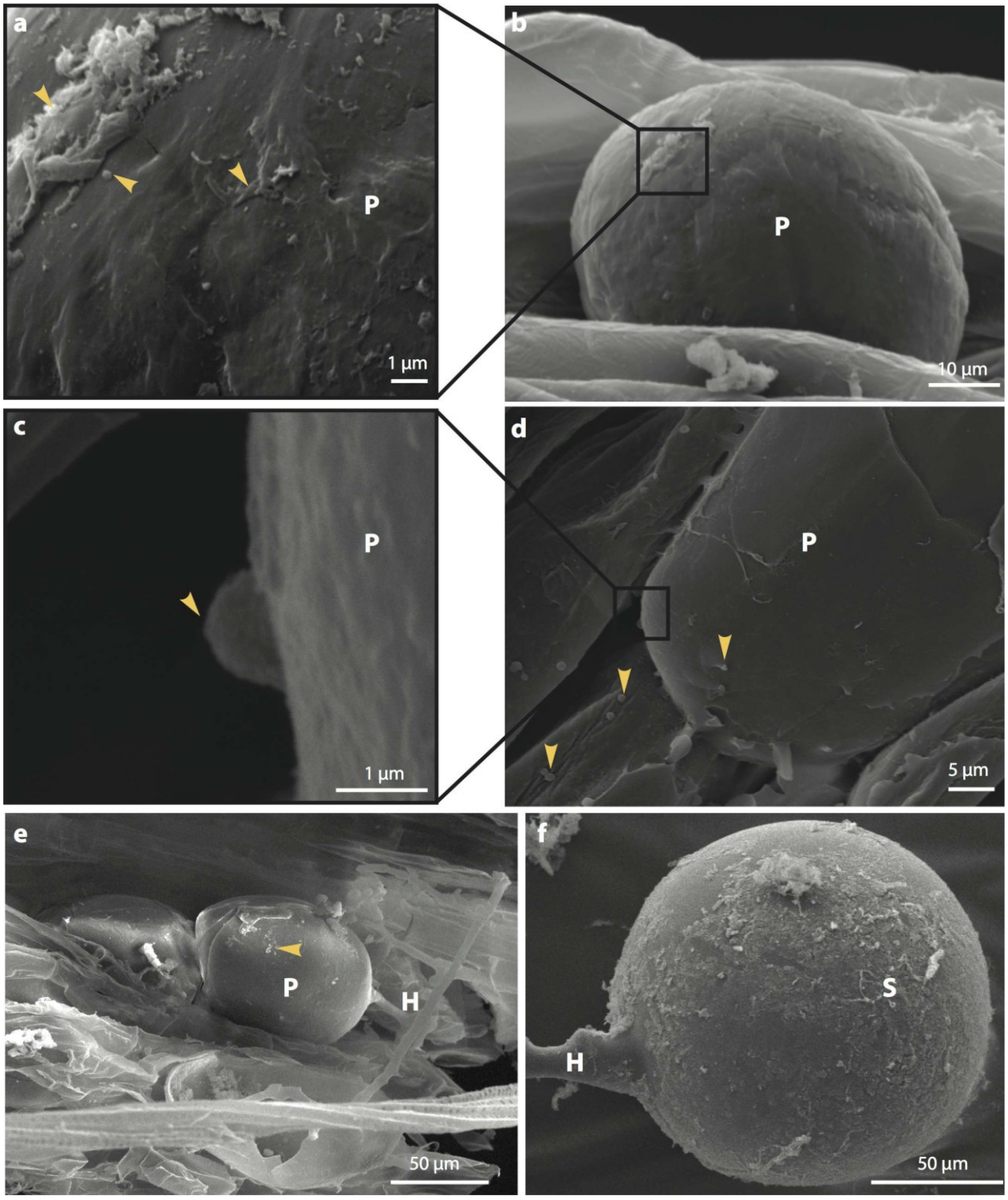


Figure 2.4. Scanning electron micrographs of dissected roots of Solidago rugosa showing AMF hyphae (H) and propagules (P) on which bacterial cells and biofilm-like structures (orange arrow) can be seen attached to their surface (panels a to e). Panels b and c are magnifications of the sections selected in the panels a and d, respectively. Panel f shows an AMF spore (S) isolated from the rhizospheric soil of S. rugosa roots sampled from the contaminated soil. Although the spore surface was washed several times with sterilized water, many microorganisms are visibly still attached to its cell wall surface.

2.6. Conclusions

In this report, we described AMF diversity and associated bacteria from individual intraradical propagules isolated from roots of plants spontaneously growing in sediments of an extreme petroleum hydrocarbon-polluted basin. Based on WGA, cloning, and sequencing on individual AMF propagules, our results showed that intraradical propagules of AMF harbor a highly diversified bacterial community. However, further investigations will be needed to determine the functions and putative roles of these bacteria in mycorrhizal symbiosis and in phytoremediation.

2.7. Acknowledgements

This project was supported by the Genome Canada and Genome Quebec funded Genorem project which are greatly acknowledged. We thank Petromont Inc. (ConocoPhillips Canada) for allowing us to access to the Varennes field site. We also thank Franck Stefani for support in sequence analyses and Karen Fisher-Favret for comments on the manuscript.

Mise en contexte chapitre 3 - Effets des niveaux de contamination et des espèces de plantes hôtes sur la biodiversité et la structure des communautés microbiennes associées aux CMA.

Dans le chapitre précédent, les techniques de microscopie, de clonage et de séquençage ont montré que même dans des conditions extrêmes, les CMA sont capables de s'associer avec plusieurs espèces de bactéries du sol. Cependant, nous ne savons pas comment les niveaux de contamination en HP et les espèces de plantes peuvent influencer la biodiversité et la structure des communautés microbiennes associées aux CMA. De plus, nous ne savons pas si les espèces de CMA peuvent aussi jouer un rôle dans la structure de ces communautés microbiennes associées à leurs spores et mycéliums ?

Dans ce chapitre, en utilisant les techniques de séquençage à haut débit, nous avons étudié la biodiversité microbienne associée aux spores des CMA extraites à partir de trois espèces de plantes qui poussent dans trois bassins de décantation ayant des niveaux de contaminations différents.

Les résultats de ce chapitre ont fait l'objet d'une publication dans la revue *Environmental Microbiology* en 2016.

Ma contribution dans cet article est la participation à la mise en place de l'expérience et la réalisation de la totalité des travaux d'échantillonnage et de laboratoire. J'ai également fait toutes les analyses bioinformatiques et statistiques. J'ai aussi participé à la rédaction et j'ai fait toutes les illustrations présentées dans l'article.

Chapitre 3 - Petroleum hydrocarbon contamination, plant identity and arbuscular mycorrhizal fungal community determine assemblages of the AMF spore-associated microbes

Bachir Iffis, Marc St-Arnaud and Mohamed Hijri

Institut de Recherche en Biologie Végétale, Département de Sciences Biologiques, Université de Montréal, 4101 Sherbrooke est, Montréal (Québec) H1X 2B2, Canada

Published in : *Environmental microbiology*, July 4, 2016. doi: 10.1111/1462-2920.13438

Author contributions : Conceived and designed the experiments: BI MH. Performed the experiments: BI. Analyzed the data: BI. Contributed reagents/materials/analysis tools: BI. MH. Wrote the paper: BI MH MSA.

3.1. Abstract

The root-associated microbiome is a key determinant of pollutant degradation, soil nutrient availability and plant biomass productivity, but could not be examined in depth prior to recent advances in high-throughput sequencing. Arbuscular mycorrhizal fungi (AMF) form symbioses with the majority of vascular plants. They are known to enhance mineral uptake and promote plant growth, and are postulated to influence the processes involved in phytoremediation. Amplicon sequencing approaches have previously shown that petroleum hydrocarbon pollutant (PHP) concentration strongly influences AMF community structure in in situ phytoremediation experiments. We examined how AMF communities and their spore-associated microbiomes were structured within the rhizosphere of three plant species growing spontaneously in three distinct waste decantation basins of a former petrochemical plant. Our results show that the AMF community was only affected by PHP concentrations, while the AMF-associated fungal and bacterial communities were significantly affected by both PHP concentrations and plant species identity. We also found that some AMF taxa were either positively or negatively correlated with some fungal and bacterial groups. Our results suggest that, in addition to PHP concentrations and plant species identity, AMF community composition may also shape the community structure of bacteria and fungi associated with AMF spores.

3.2. Keywords

Arbuscular mycorrhizal fungi; Petroleum hydrocarbon pollutants; 454 high throughput sequencing; Spore-associated microbes.

3.3. Introduction

Arbuscular mycorrhizal fungi (AMF) constitute a widespread soil fungal group which establishes symbiotic associations with the majority of vascular plants (Smith & Read, 2008). Once AMF colonize plant roots, they develop an extraradical hyphal network, promote nutrient uptake, in particular phosphorus and nitrogen, and protect plants against soil-borne pathogens (St-Arnaud & Vujanovic, 2007; Smith & Read, 2008; Ismail *et al.*, 2011). In the last few years, several studies showed that AMF can also influence the fate of trace elements and petroleum hydrocarbon pollutants (PHP) in soils (Gao *et al.*, 2010; Gao *et al.*, 2011b; Hassan *et al.*, 2013). As fungi that extend their hypha into the rhizosphere and the surrounding soil, AMF not only interact with plants, but also with a myriad of soil microbes. Several *in vivo* and *in vitro* studies have demonstrated that a large range of bacterial species are living on the surface and/or inside mycelia, spores and intraradical propagules of AMF (Bianciotto *et al.*, 1996; Scheublin *et al.*, 2010; Lecomte *et al.*, 2011; Cruz & Ishii, 2012; Iffis *et al.*, 2014; Agnolucci *et al.*, 2015). Some fungal taxa belonging to *Ascomycota* and *Chytridiomycota* have also been reported to be associated with AMF mycelia (Paulitz & Menge, 1984; Hijri *et al.*, 2002; Iffis *et al.*, 2014).

Even if there is no clear evidence that AMF directly metabolize and degrade petroleum hydrocarbons, they might stimulate soil microbial metabolic activity by releasing nutrients in the mycosphere similar to plant root exudates in the rhizosphere (Barea *et al.*, 2005; Boer *et al.*, 2005; Frey-Klett *et al.*, 2007; Bonfante & Anca, 2009). This may lead to an acceleration of degradation of organic pollutants by mycosphere microbes as well as to the translocation and immobilization of trace elements (Joner *et al.*, 2001; Liu & Dalpé, 2009; Hernández-Ortega *et*

al., 2012; Hassan *et al.*, 2013), and their accumulation in plant tissues (Jing *et al.*, 2007). Previous reports have shown that AMF can improve soil resilience to organic pollution (Joner *et al.*, 2001; Leyval *et al.*, 2002; Joner & Leyval, 2003; Liu *et al.*, 2004; Volante *et al.*, 2005). However, these studies were mostly done only under greenhouse or *in vitro* conditions. Hassan *et al.* (2014) reported that PHP contaminant concentrations strongly influenced the AMF community structure within the rhizosphere of 11 willow cultivars planted across PHP contaminated soils at the site of a former petrochemical plant. These authors also showed that different AMF families dominated at each contaminant level. Determination of the best approach for selecting competent microbes for bioremediation should be based on prior knowledge of the microbial communities inhabiting the target site (Tyagi *et al.*, 2011), since long-term exposure to a contaminant may have allowed different microbes to develop tolerance to highly polluted conditions or favored the proliferation of taxa able to metabolize or accumulate the pollutants.

In this study, we report the effects of PHP concentrations and host plant identity on the diversity of AMF and of their spore-associated microbiomes in three distinct waste decantation basins located on the site of a former petrochemical plant. AMF associated microbiome refers here to the whole community of bacteria and fungi living on the surface and/or inside AMF spores. We further investigated the relationships between community structure of AMF and of their spore-associated microbial communities.

We sampled rhizospheric soils and roots of three plant species, *Solidago canadensis*, *Populus balsamifera* and *Lycopus europaeus* growing spontaneously in the three contaminated basins, leading to 27 samples. From each rhizospheric soil sample, we manually collected one

thousand AMF spores using wet sieving and sucrose gradient centrifugation. The rationale for choosing spores relies on the fact that they can easily be discriminated from those of other fungi. Total DNA of spore samples was individually extracted, PCR-amplified and sequenced using the Roche GS-FLX+ platform targeting the 18S rRNA gene for AMF, ITS for fungi and 16S rRNA gene for bacteria. Our results showed that the community structures of AMF were only affected by PHP concentration, while the AMF-associated microbes were significantly affected by both PHP concentration and plant species identity. The proportions of taxa belonging to known petroleum-degrading microorganisms were higher in the highly contaminated basin than in the lowest contaminated basin. We also observed that in addition to the PHP contaminants and plant species effects, AMF may also shape the microbial communities associated with their spores. This investigation is the most comprehensive study on an AMF spore-associated microbiome, whose diversity is found to be much more complex than previously shown.

3.4. Materials and methods

3.4.1. Experimental design and sampling

Field sampling was carried out on October 18th, 2013, in three decantation basins in which petroleum hydrocarbon wastes have been dumped by a former petrochemical plant located on the south shore of the St-Lawrence River near Montreal, QC, Canada (45.70 N, 73.43 W). More details and description of this site can be found in Bell *et al.* (2014). An exhaustive inventory of spontaneous vegetation growing on this site was also reported by Desjardins *et al.* (2014), who identified 23 plant species. We selected *Solidago canadensis*, *Populus balsamifera* and *Lycopus europaeus* for this study based on their abundance, their

presence in each of the three adjacent basins and their ability to form mycorrhizal symbiosis. In each basin, we collected three individual plants of each plant species with the soil surrounding their roots, totaling 27 plant samples (three basins × three plant species × three replicates). The fresh soil attached to each plant root system was collected and the AMF spores were separately extracted from each sample as described below. To confirm the AMF colonization of the sampled plant roots, a subsample of each root system was cleaned under tap water, cut into 1 cm long pieces, stained using 1 % Trypan blue and observed under the microscope (Giovannetti & Mosse, 1980).

3.4.2. Soil chemical analysis

Nine composite soil samples, each formed from a mix of the soil from the three replicates of each plant species per basin, were sent for polycyclic aromatic hydrocarbons (PAHs) and total alkanes (C10–C50) analyses to a commercial laboratory service (Maxxam, Montreal, QC). Soil chemical analyses are summarized in Supporting Information Table S3.10. Basin 1 showed the highest contamination level (termed as HC), followed by Basin 3, with a moderate contamination level (termed as MC), while Basin 2 showed the lowest contamination level, with concentrations comparable to a non-contaminated soil (termed as LC), according to the definitions used by the government of the province of Québec, Canada (http://www.mddelcc.gouv.qc.ca).

3.4.3. AMF spore harvesting

AMF spores were separately extracted from the fresh soil of each individual plant sample. AMF spores were extracted using wet sieving followed by sucrose gradient centrifugation according to Walker *et al.* (1982), with some minor modifications. Briefly, the

soil surrounding roots and a part of the bulk soils were collected by shaking and scratching plants roots in sterile plastic bags. Then, the collected soil was homogenized and a 50 g subsample was mixed in a 1 L container half full of tap water. The soil suspension was passed through a series of five superposed sieves with mesh sizes of 1 mm (top sieve), 500 µm, 250 µm, 150 µm, and 40 µm (bottom sieve). The sieve assemblage was mounted in a vibratory sieve shaker apparatus (AS 200 basic, Retsch), operated at 30 Hz for 3 min under a mild water jet. The soil particles collected in 250 µm, 500 µm and 1 mm sieves were discarded, while those collected in 40 µm and 150 µm sieves were put into 50 ml tubes containing 20 ml of water. The suspensions were then subjected to sucrose gradient centrifugation and, for each sample, approximately 1000 spores were handpicked under a stereomicroscope using a 100 µl micropipette and collected in 1.5 ml microtubes. DNA extractions were performed directly on the fresh spores as described below.

An additional 50 spores sample was harvested from each plant and they were subjected to electron scanning microscopy using a Quanta 200 3D SEM (FEI, Burlington, MA), as described in Iffis *et al.* (2014).

3.4.4. DNA extraction

Spores were washed three times with sterilized distilled water in order to remove the soil particles adhering on spore surfaces. DNA extraction was then performed using the DNeasy Plant Mini Kit (QIAGEN, Toronto, ON, Canada) following the manufacturer's instructions. DNA was eluted in 50 µl of elution buffer, and quantified using a Qubit 2.0 fluorometer (Life Technologies, Burlington, ON, Canada). DNA concentrations ranged between 0.4 and 10 ng/µl. The extracted DNA samples were stored at -20 °C until use.

3.4.5. Polymerase chain reactions

To identify the AMF taxa, the extracted DNA samples were subjected to PCR amplifications targeting a partial 18S rRNA gene fragment with the primer pair AM1 and NS31, which produce a fragment of approximately 550 bp (Simon *et al.*, 1992; Lee *et al.*, 2008). It is known that AM1 is not meant to amplify *Archaeosporales* and *Paraglomerales* (Lee *et al.*, 2008; Van Geel *et al.*, 2014), so our results might have underestimated these taxa if present (although some singletons of *Paraglomus* were produced). We initially used the AML1-AML2 primer pair aiming to amplify a more diverse set of Glomeromycota taxa but the sequence quality was bad and for this reason we used AM1 and NS31 primers. To identify the AMF spore-associated microbes and assess their community structure, two other PCRs were performed with primers UnivBactF 9 and BSR534/18, and ITS1F and ITS4, targeting the 16S rRNA gene of bacteria and the ITS regions of fungi, respectively (Bell *et al.*, 2014). All primers cited above were tagged with adaptors and unique multiplex identifier (MID) tags from the extended MID set recommended by Roche Diagnostics (Roche, 2009).

PCRs were performed in 50 µl volumes containing 5 µl of 10× PCR buffer, 0.2 mM of dNTP mix, 1 µl of BSA (100 mg/ml), 1 µl $MgCl_2$ (25 mM), 0.4 mM of each primer, 2 µl of DNA template and 1 U of *Taq* DNA polymerase (QIAGEN, Toronto, ON). PCRs were run on a thermal cycler Pro S thermocycler (Eppendorf, Mississauga, ON, Canada) using an initial denaturation at 95 °C for 5 min followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min, and a final elongation step at 72 °C for 10 min. After electrophoresis separation and UV light visualization, PCRs products were purified with the QIAquick Gel Extraction Kit (QIAGEN, Toronto, ON) following the manufacturer's instructions. DNA concentrations of the purified PCR products were measured using the Qubit 2.0 fluorometer and three pools

were prepared by mixing equimolar volume of PCR products. The first pool contained 18S rRNA gene amplicons, the second contained 16S rRNA gene amplicons and the third contained ITS amplicons. These pools were sent for sequencing to the Genome Quebec Innovation Centre using the Roche 454 FLX+ pyrosequencing platform (Roche, Branford, CT, USA). One eighth of each sequencing plate was used for sequencing each of the 18S rRNA gene, 16S rRNA gene and ITS amplicons.

3.4.6. Bioinformatic processing

The 454 datasets were processed in Mothur (v.1.34.4) (Schloss *et al.*, 2009). The partial 16S rRNA gene and ITS sequences were processed as described in Mothur wiki (http://www.mothur.org/) and in Bell *et al.* (2014) with some minor changes. For the bacterial 16S sequences, the "*minlength*" parameter of the command "*trim.seqs*" was adjusted to 300 instead of 200. The ITS sequences were processed three times: (i) on the overall dataset, including AMF sequences; (ii) on a subset containing only the ITS sequences belonging to AMF; (iii) on a subset containing only the non-AMF fungal ITS sequences. The quality filtering steps were done similarly between the three ITS datasets, whereas the OTUs grouping, taxonomy and standardizing were performed separately on the different ITS datasets.

The AMF 18S rRNA gene sequences were analyzed as described in Hassan *et al.* (2014) with the following modifications: "*minlength*" and "*qwindowaverage*" of "*trim.seqs*" command were 350 and 35 respectively. Parameters of "*screen.seqs*" command were: end=26790, optimize=start, criteria=90 and processors=2. After removing non-AMF sequences and standardizing the dataset using subsampling in order to have an equal number of reads per sample, the 18S rRNA gene sequences were assigned to OTU groups at a 97 % of

sequence identity using Geneious version 6 (Biomatterts Limited, Auckland, New Zealand), then a Blast search was carried out in the MaarjAM database (Öpik *et al.*, 2010).

In total, five sequence datasets were obtained: AMF 18S rRNA genes, AMF ITS, non-AMF ITS, overall fungal ITS and bacterial 16S rRNA genes. For each dataset, reads were assigned to OTUs at 97 % of sequence identity. Singletons of each dataset were removed (OTUs that were globally represented by less than 2 reads in each dataset). Datasets were randomly standardized by subsampling to the sample size that showed the lowest number of reads.

All 454 databases generated in this study were deposited in the NCBI Sequence Read Archive and are available under the project number SRP069084.

OTUs tables of each dataset, after and before subsampling, are provided as supplemental tables in the excel file named list of the OTUs tables (Table S3.1 to Table S3.4, Table S3.6 and Table S3.11 to Table S3.13).

3.4.7. Statistical analysis

All statistical analyses were performed in R software (version 3.1.1). After verification of the normalities by Shapiro–Wilk test, the effect of contamination concentration and plant species identity on the Shannon diversity index of the microbial communities, which are normally distributed, were tested by ANOVA using the "Rcmdr" package. To verify the efficiency of our sampling efforts and sequencing depth, rarefaction curves were drawn for each individual sample using the "rarefy" function in the "vegan" package. To test the effect of the different factors on the community structures, PERMANOVA analyses were performed using the "adonis" function in the "vegan" package on Bray–Curtis values obtained from the

community structures matrices previously normalized by the Hellinger transformation. To test the homogeneity of dispersion of the different communities (AMF, other fungi and bacteria) against PHP concentration and plant species identity, beta-dispersion analyses were performed on the Bray–Curtis matrices using the "betadisper" function in the "vegan" package. Principal coordinate analyses (PCoA) were performed in order to visualize the effects of contamination levels and plant species identity on community composition. PCoA ordinations were calculated using the "cmdscale" function on Bray–Curtis values. In order to reveal which fungal or bacterial groups were affected by the contamination level or plant species, Kruskal-Wallis tests were performed on the most abundant: five OTUs of AMF 18S rRNA gene dataset, ten OTUs of the AMF ITS dataset, twenty OTUs of the non-AMF ITS dataset and fifty OTUs of the bacterial 16S rRNA gene dataset. Kruskal–Wallis tests were also performed on the different taxonomic levels of the different datasets. Relative abundances were calculated with Excel on the different datasets to visualize the percentage of the taxonomic affiliations across sites and plant species. Co-inertia analysis (CoIA) was carried out in order to display the relationships between: (i) the community structures of AMF 18S rRNA gene (at the genera level) and bacterial 16S rRNA gene (at the class level), (ii) the community structures of AMF 18S rRNA gene (genera) and non-AMF fungi (classes). Co-inertia analysis was performed with the package "ade4" following the protocol described in Legendre and Legendre (2012). CoIA considered as an alternative method to the canonical correlation analysis with the main function based on the matching of two datasets and projection of their variables in the same space (named co-inertia plane) in order to visualize their co-variations (relationships). One advantage of this method is that it allows to study the co-variation between two response matrices (symmetric model), contrarily to some canonical analysis (e.g.

redundancy analysis) which requires a response matrix and an explanatory matrix (Dolédec & Chessel, 1994; Dray *et al.*, 2003; Legendre & Legendre, 2012). Here, we used the CoIA approach to study the relationships between AMF and AMF spore-associated bacterial and fungal communities because the three datasets (AMF, fungi, and bacteria) were produced from the same samples (the AMF spores), and therefore no dataset would represent the response matrix or the explanatory matrix. To support the Co-inertia analysis, Spearman correlation tests were performed using the "corr.test" function in the "psych" package in order to calculate the different paired correlations between AMF genera and the microbe classes (bacteria and fungi classes). *P*-values and corrected *P*-values (using false discovery rate method) were also generated with the "corr.test" function for each correlation coefficient. The array correlation matrices were drawn using the "levelplot" functions in the "lattice" package. A variance partitioning analysis was performed in order to test the contribution of AMF communities, PHP concentrations and plant species identity on the variation of bacterial and fungal community structures using the "varpart" function in "vegan" package in R (more details on variance partitioning analysis are available in Supplemental Material).

3.5. Results

3.5.1. Microscopy observations

Root staining and light microscopy clearly confirmed the mycorrhizal colonization of the roots of the three selected plant species. Typical vesicles and intraradical hyphae of AMF were clearly visible in all root fragments examined (Figure S3.1.). Scanning electron microscopy of AMF spores showed the presence of aggregates and biofilm-like structures adhering to the surface of spores, whose morphology varied greatly among examined spores.

The presence of thin structure, typical of fungal filaments, was also observed on the surface of some spores (Figure S3.2.).

3.5.2. OTUs affiliations of the AMF spores and of their associated-microbiomes

After quality filtering and standardizing the number of sequences of the different datasets, the AMF 18S rRNA gene dataset allowed us to retrieve a total of 21465 sequences (795 per sample across the 27 samples) that were assigned to 27 OTUs (Table S3.1). For the whole fungal ITS dataset, 16983 sequences (629 per sample) were obtained and assigned to 141 OTUs. In the AMF ITS subdataset, we obtained a total of 6507 sequences (241 per sample) that were assigned to 48 OTUs (Table S3.2). For the fungal ITS subdataset (excluding the AMF sequences), a total of 1269 sequences (47 per sample) were obtained and they were assigned to 66 fungal OTUs (Table S3.3). A total of 27324 sequences (1012 sequences per sample) were retrieved from the 16S rRNA gene dataset after bioinformatic processing, and were assigned to 1080 bacterial OTUs (Table S3.4). For each AMF, fungi and bacteria datasets, rarefaction curves showed that the sampling efforts were close to the saturation for the all samples, and the Good's coverage values were ranked between 0.85 and 1 (Figure S3.3. and Table S3.5.).

Blast searches of the AMF 18S rRNA gene sequences in MaarjAM database (http://maarjam.botany.ut.ee/) (Öpik *et al.*, 2010) allowed us to assign the OTUs to virtual taxa (VTX) which were mainly represented by AMF taxa belonging to the genera *Diversispora* (46%), *Glomus* (28 %), *Acaulospora* (16 %) and *Claroideoglomus* (10 %) (Table S3.5. and Figure S3.4. A). Other taxa belonging to the genus *Paraglomus* were also found as singletons but were removed after dataset filtering (Table S3.6). Interestingly, the AMF taxa profile

obtained from the ITS dataset was comparable in term of abundance for the most dominant taxa identified in the AMF 18S rRNA gene dataset, with genera *Diversispora* and *Glomus* forming 43 % and 27 % of the reads, respectively (Figure S3.4. B). However, in contrast to the 18S rRNA gene dataset, the proportions of *Acaulospora* and *Claroideoglomus* were lower in the AMF ITS dataset. Furthermore, the AMF ITS dataset showed the presence of other genera such as *Rhizophagus*, *Entrophospora*, *Funneliformis*, as well as unclassified *Archaeosporaceae*, although most of them were found at low abundance.

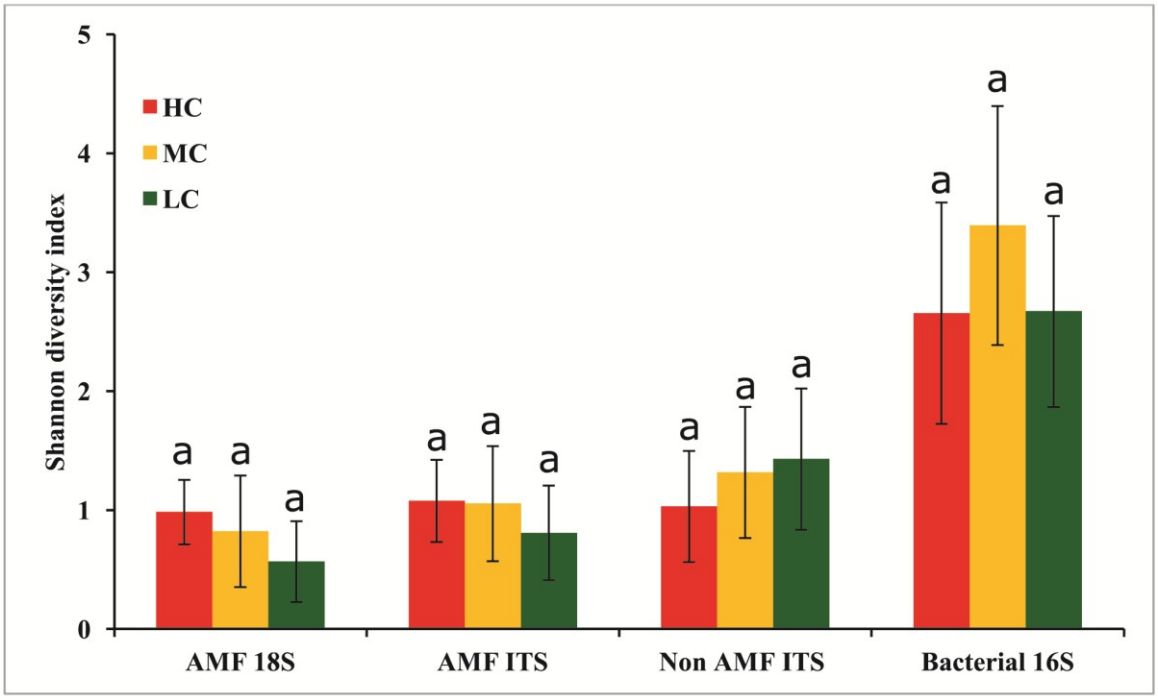
Because the total DNA was directly extracted from washed AMF spores, it was not unexpected that the AMF sequences represented 68 % of the reads in the whole fungal ITS dataset (Figure S3.4. D). The AMF sequences were therefore sorted out to focus on the non-AMF fungal diversity associated with the spores. After sorting, 55 % of the remaining ITS dataset was composed of *Pezizomycetes* (13 %) and *Dothideomycetes* (13 %) followed by *Chytridiomycetes*, *Sordariomycetes* and *Microbotryomycetes* at a proportion of 9 % each (Figure S3.4. C). The other fungal classes of *Agaricomycetes*, *Leotiomycetes*, *Tremellomycetes* and *Ustilaginomycetes* were also found but their proportions were lower than 1 %. The remaining 45 % of fungal sequences were assigned to unclassified fungi.

The bacterial 16S rRNA gene dataset profile was composed of taxa belonging to 33 bacterial classes. The most abundant classes were *Gammaproteobacteria* (49 %), *Betaproteobacteria* (23 %), *Actinobacteria* (11 %) and *Alphaproteobacteria* (6 %). The other bacterial classes were present at proportion lower than 3 % (Figure S3.4. E).

3.5.3. Effect of PHP concentrations and plant species identity on the diversity and community structure of AMF

ANOVA tests revealed that Shannon diversity indices of the 18S rRNA gene dataset were near-significantly affected by contamination concentration (P = 0.073), with a highest diversity in the HC site and lowest diversity in LC site. Whereas, Shannon diversity indices of the ITS dataset were not affected by both contamination concentration and plant species identity (Figure 3.1. and Table 3.1). However, PERMANOVA analysis revealed a significant effect of contamination concentration on the community structure of both AMF 18S rRNA gene and AMF ITS sequences (P = 0.004 and 0.016, respectively). Beta-dispersion analyses showed a near-significant dispersion of AMF 18S rRNA gene OTUs and a significant dispersion of AMF ITS OTUs across contamination concentrations (P = 0.093 and 0.042, respectively) (Table 3.1). A post-hoc Tukey's HSD test showed that the AMF ITS communities were more variable in the HC site and MC site than in the LC site (Figure S3.5. D). The PCoA ordinations showed a clear grouping of AMF community structure in response to the contamination level, in particular between HC and LC sites (Figure 3.2. A and B).

(A)



(B)

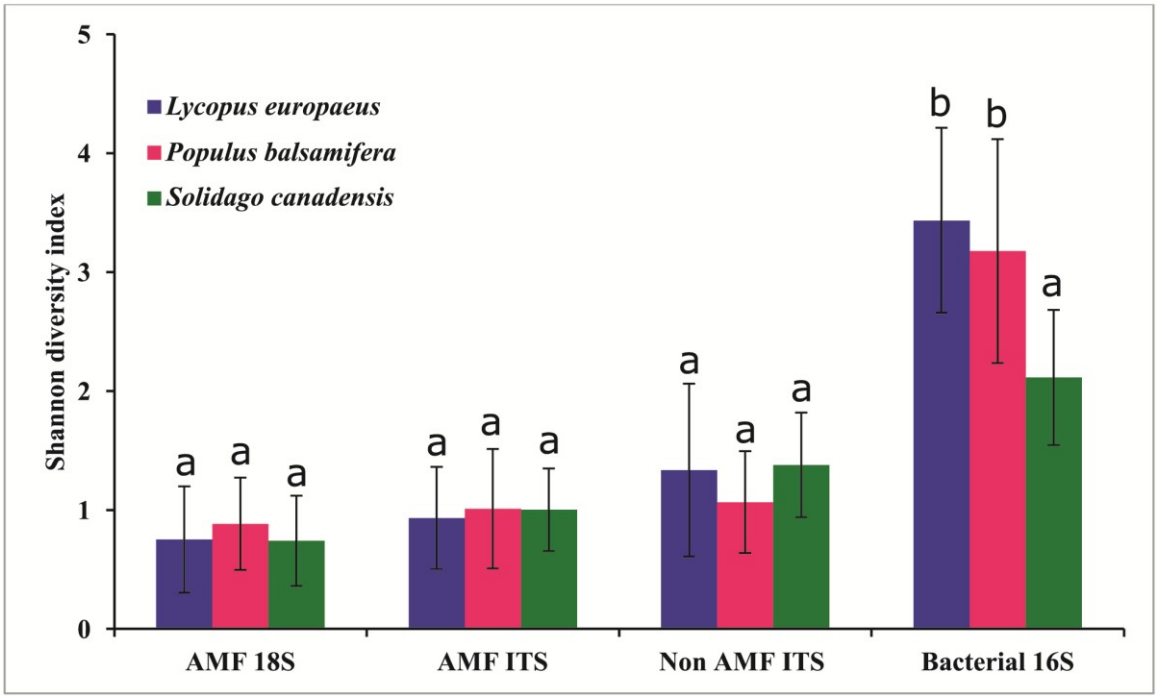


Figure 3.1. Comparison of Shannon diversity indices of the different pyrosequencing datasets across contamination concentrations (A), and plant species (B).
Within each microbial group, means with the same letter are not significantly different by a Tukey's range test ($P < 0.05$). HC = high contamination, MC = moderate contamination, LC = low contamination.

The Kruskal-Wallis tests performed on most abundant OTUs of the AMF 18S rRNA gene dataset revealed that the contamination concentration significantly affected the OTUs belonging to *Diversispora* sp. VTX00061 (P = 0.002), *Rhizophagus* sp. VTX00113 (P = 0.003), and *Claroideoglomus* sp. *VTX00193* (P = 0.021) (Table S3.7. A). Relative abundances showed that the genus *Diversispora* was the most abundant taxon at the LC site (78.07 %) and the MC site (48.23 %), (Figure 3.3 A). By contrast, genera *Glomus* and *Claroideoglomus* were mostly represented in the HC site with proportion of 42.43 % and 17.49 %, respectively, and the MC site (29.51 % and 11 %, respectively) (Figure 3.3 A). The Kruskal-Wallis tests and the relative abundances calculated from the AMF ITS dataset were similar to those performed on the 18S rRNA gene dataset. The only notable difference was that, in addition to genera affected by the contamination concentration in 18S rRNA gene dataset, the genus *Rhizophagus* was also significantly (P < 0.001) affected by the PHP contamination concentration in the AMF ITS dataset, with a high abundance in the HC site (39.37 %) and MC site (16.96 %) (Figure 3.3 B; Table S3.7. A and B). However, no effect of plant species identity was noted on AMF community structure for either the 18S rRNA gene or the ITS sequence datasets.

Table 3.1. ANOVA on Shannon diversity indices, PERMANOVA and Beta-dispersion analyses on the community structures of the different pyrosequencing datasets (n = 9 for plants species and contamination levels, and n = 3 for the interaction effects). The Bolded values are significant at $P < 0.05$. NA mean not calculated.

	Contamination level			Plant species			Interaction effects		
	ANOVA	PERMANOVA	Beta-dispersion	ANOVA	PERMANOVA	Beta-dispersion	ANOVA	PERMANOVA	Beta-dispersion
AMF 18S	0.073	**0.004**	0.093	0.711	0.898	0.876	0.180	0.241	NA
AMF ITS	0.326	**0.016**	**0.042**	0.917	0.716	0.887	0.168	**0.023**	NA
ITS (excluding AMF)	0.292	**0.002**	0.77	0.439	**0.010**	0.585	0.379	0.354	NA
Bacterial 16S	0.171	**0.010**	0.383	**0.00336**	**0.001**	0.088	0.960	0.186	NA

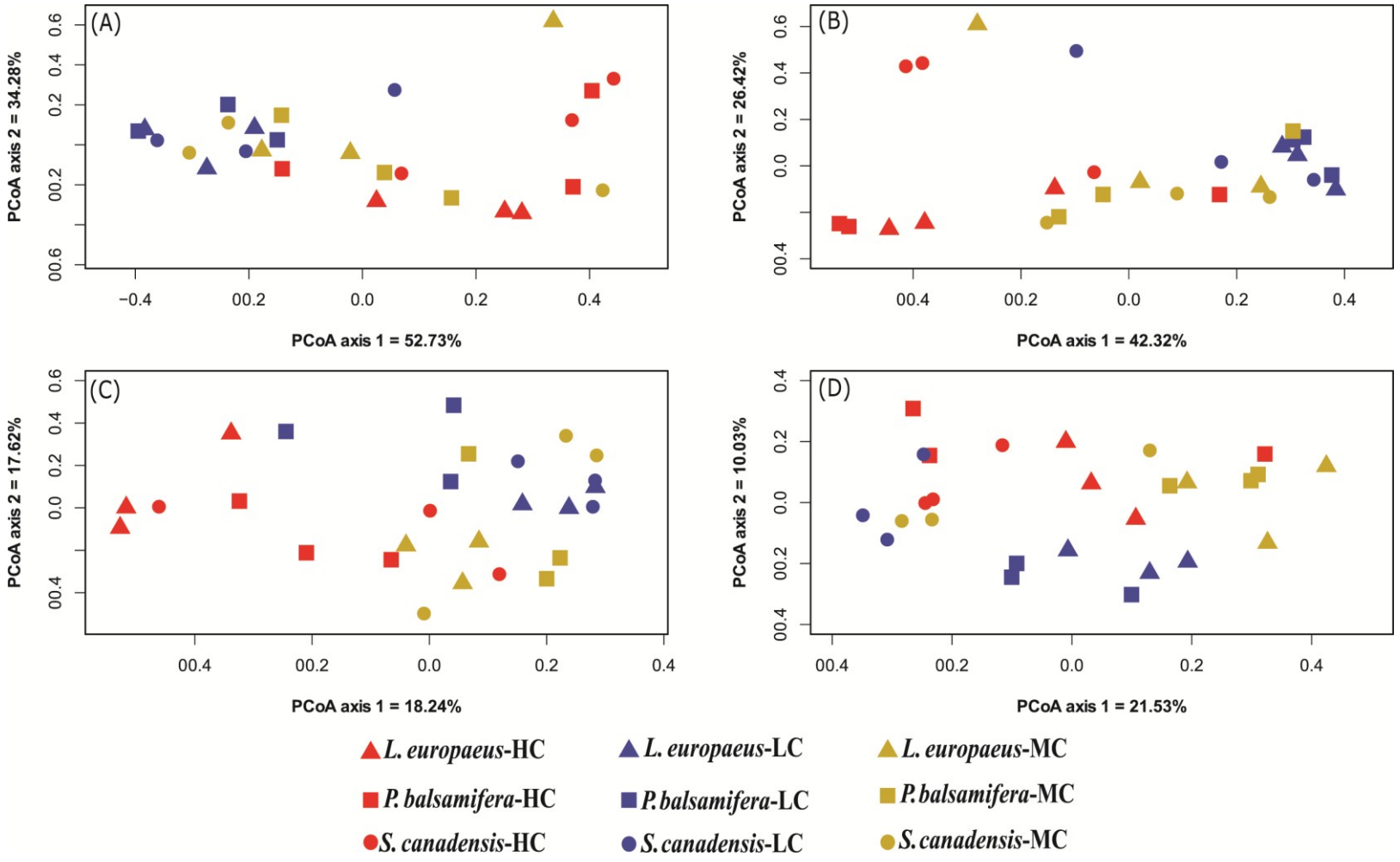


Figure 3.2. Principal coordinates analysis (PCoA) showing the community compositions assignments of: (A) AMF 18S rRNA gene, (B) AMF ITS, (C) non AMF ITS, and (D) 16S rRNA gene across contamination level and plant species.
PERMANOVA analysis showed significant effects of the contamination concentrations on the community composition of AMF ITS, fungal ITS and 16S rRNA gene (n = 9, P = 0.016, 0.002 and 0.010, respectively). Significant effects of plant species were also observed on the community compositions of non AMF ITS and 16S rRNA gene (n = 9, P = 0.010 and 0.001, respectively).

3.5.1. Effect of PHP concentration and plant species identity on the diversity and community structure of AMF spore-associated fungi

ANOVA tests showed that there was no effect of PHP contamination concentration and plant species identity on Shannon diversity indices of fungi, excluding the AMF (Figure 3.1. and Table 3.1). However, PERMANOVA tests revealed that the community structure of fungi was significantly affected by both contamination concentration ($P = 0.002$) and plant species identity ($P = 0.01$), while Beta-dispersion analyses showed that the AMF spore-associated fungi were homogeneously dispersed across plant species identity and PHP contaminated sites (Table 3.1; Figure S3.5. E and F). The PCoA plot showed a change in community structure of fungi across contamination concentrations and distinct groupings were observed between HC and LC sites (Figure 3.2.C).

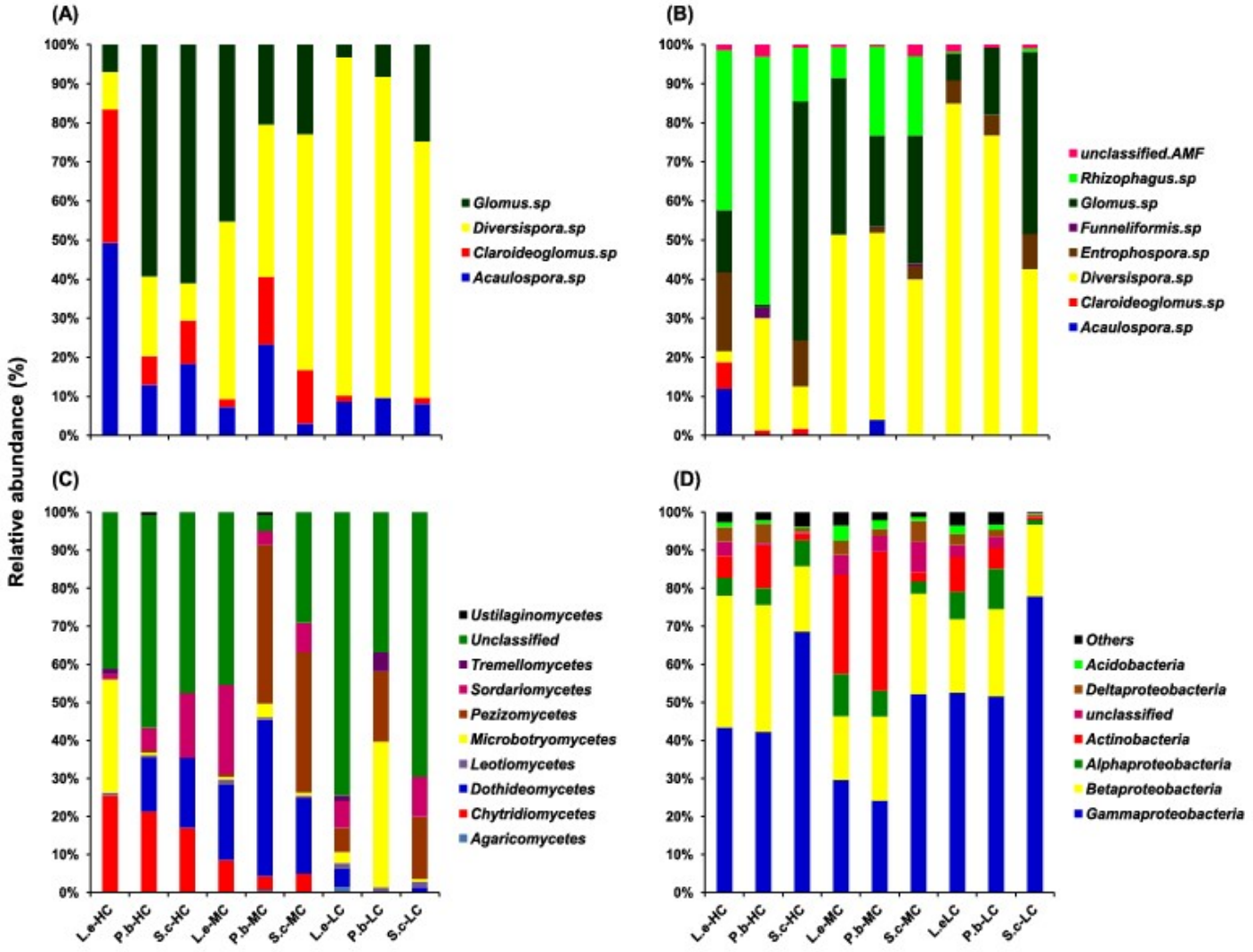


Figure 3.3. Relative abundances of major: (A) genera of AMF 18S rRNA gene dataset, (B) genera of AMF ITS dataset, (C) classes of fungal ITS dataset, and (D) bacterial classes of 16S rRNA gene dataset.

So-HC: *S. canadensis* High Contamination; Po-HC: *P. balsamifera* High Contamination; Ly-HC: *L. europaeus* High Contamination; So-MC: *S. canadensis* Moderate Contamination; Po-MC: *P. balsamifera* Moderate Contamination; Ly-MC: *L. europaeus* Moderate Contamination; So-LC: *S. canadensis* Low Contamination; Po-LC: *P. balsamifera* Low Contamination; Ly-LC: *L. europaeus* Low Contamination.

When comparing the community composition of fungi across contamination concentrations, Kruskal-Wallis tests revealed that taxa *Pulvinula constellatio* (*Pezizomycetes*) and *Spizellomyces plurigibbosus* (*Chytridiomycetes*) were significantly affected by PHP concentration (P = 0.008 and 0.015, respectively) (Table S3.7. C). Kruskal-Wallis tests were performed also at the fungal class level and allowed us to note a significant effect of the contamination concentration on the proportions of *Pezizomycetes* and *Chytridiomycetes* ($P <$ 0.05), and a near-significant effect on *Dothideomycetes* ($P < 0.1$) (data not showed).

Relative abundances showed that *Pezizomycetes* was the most dominant fungal class in the MC (26.74 %) and LC (13.71 %) sites, in particular with *Populus balsamifera* and *Solidago Canadensis*. *Dothideomycetes* was higher in MC (26.95 %) and HC (10.87 %) sites compared to the LC site (2.12 %). *Chytridiomycetes* which were represented mainly by the family of *Spizellomycetaceae*, were observed only in HC (21.27 %) and MC sites (5.67 %). In contrasting, no *Chytridiomycetes* OTU was detected in the LC site (Figure 3.3 C and Figure 3.4 C).

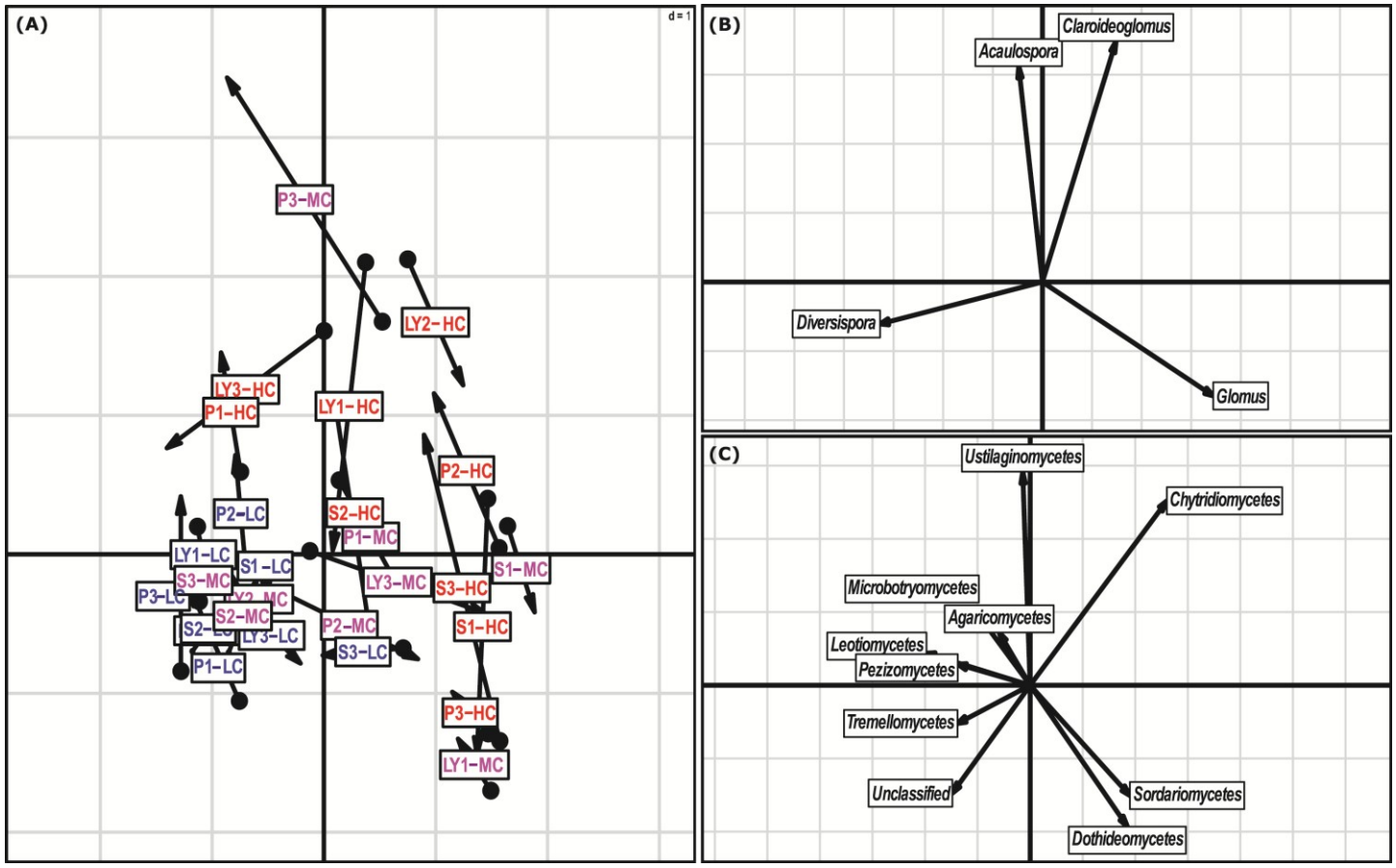


Figure 3.4. Co-inertia analysis showing the relationship between AMF genera obtained from 18S rRNA gene dataset and fungal classes (*RV* = 0.29, *P* = 0.004). A) Projection of both AMF genera and AMF-associated fungal classes onto co-inertia plane. B) Projection of AMF genera onto co-inertia plane. C) Projection of AMF-associated fungal classes onto co-inertia plane.

The base of each arrow (black circle) represents AMF genera, while the tip represents fungal classes. The lengths of the arrows indicate the degree of concordance between AMF genera and their associated microorganisms. The best concordances were observed in the shortest arrows length.

Kruskal-Wallis tests performed on the most abundant fungal OTUs across plant species showed a slightly significant effect on *Leptosphaeria* sp. (Dothideomycetes) (P = 0.09), with a higher proportion in *P. balsamifera* (11.58 %) and *S. canadensis* (4.49 %) (Table S3.7. C). Kruskal-Wallis tests performed at the class level revealed that there was no significant effect of plant species identity, except *Sordariomycetes* which their proportions showed a trend toward significance across plant species. (P = 0.10). The relative abundance of *Sordariomycetes* was higher in *L. europaeus* (10.63 %) and *S. canadensis* (11.82 %) than in *P. balsamifera* (3.07 %) (Figure 3.3C).

3.5.2. Effect of PHP concentration and plant species identity on the diversity and community structure of AMF spore-associated bacteria

In contrast to AMF and other fungi, ANOVA test revealed that there was a highly significant effect of plant species identity (P = 0.003) on Shannon diversity indices of the bacterial 16S rRNA gene dataset. Tukey's range test showed that the bacterial diversity was higher in samples from *L. europaeus* and *P. balsamifera* plants compared to *S. canadensis*. There was no significant effect of PHP concentration on Shannon diversity indices of bacteria.

The PERMANOVA analysis showed that the AMF spore-associated bacterial community structure was affected both by PHP contaminant concentration (P = 0.01) and plant species identity (P = 0.001) (Figure 3.1. and Table 3.1). However, Beta-dispersion analyses showed a near-significant dispersion of AMF spore-associated bacterial communities across plant species identity and a homogeneous dispersion across contamination concentration (P = 0.088 and 0.383, respectively). Boxplot of beta-dispersion per plant species identity showed that the AMF spore-associated bacterial communities were more variable in *L. europaeus* and *P. balsamifera* than in *S. canadensis* (Figure S3.5. G). The PCoA ordination

performed on bacterial OTUs also showed that, in addition to the grouping across the sites, the bacterial communities were also grouped across plant species. *S. canadensis* formed a distinct group compared to *L. europaeus* and *P. balsamifera* (Figure 3.2. D and F; Figure S3.6. D). Similar PCoA plots were also obtained at the bacterial family and class levels (data not showed).

A Kruskal–Wallis test revealed that among the most abundant 50 OTUs, 35 were significantly affected either by contamination concentration (15 OTUs), plant species identity (13 OTUs), or by both (six OTUs). Most of the OTUs affected by the contamination concentration and/or plant identity belonged to the bacterial classes *Actinobacteria*, *Alphaproteobacteria, Betaproteobacteria*, *Gammaproteobacteria, Deltaproteobacteria* and *Flavobacteriia* (Table S3.7. D).

Histograms of relative abundances of the different bacterial classes across PHP contamination concentrations showed that *Gammaproteobacteria* ($P = 0.031$) were the most abundant in the LC site (60.69 %) followed by HC (51.43 %) and MC (35.35 %) sites. *Actinobacteria* ($P = 0.086$) were more abundant in MC site (21.72 %) than in HC (6.29 %) and LC (5.12 %) sites. For *Betaproteobacteria*, there was no significant effect ($P = 0.25$), but a slightly higher abundance was noted in the HC site (28.35 %) compared to the MC (21.72 %) and LC (20.36 %) sites. No significant effect was observed on *Alphaproteobacteria* ($P = 0.48$). However, a slightly significant effect of contamination concentration was measured on *Caulobacteraceae* ($P = 0.07$), which was the most abundant family of *Alphaproteobacteria*, with a higher abundance in the LC (2.76 %) and MC (1.12 %) sites (Figure 3.3D).

Comparison of the relative abundances across plant species showed that *Gammaproteobacteria* (P = 0.007) were more abundant in *S. canadensis* samples (66.23 %) than with *L. europaeus* (41.85 %) and *P. balsamifera* (39.39 %). However, *Actinobacteria* (P < 0.001)*, Deltaproteobacteria* (P = 0.026) and *Acidobacteria* groups (P < 0.001) were clearly more abundant with *L. europaeus* (13.68 %, 3.55 % and 2.51 %, respectively) and *P. balsamifera* (17.8 %, 2.9 % and 1.5 %, respectively) than with *S. canadensis* (1.65 %, 2.29 % and 0.55 %, respectively).

For *Alphaproteobacteria*, a slightly significant effect of plant identity was measured (P = 0.10), with a highest abundance in *L. europaeus* (7.69 %) and *P. balsamifera* (7.30 %) samples than with *S. canadensis* (3.78 %) (Figure 3.3 D).

3.5.3. Relationship between AMF spores and their associated fungi and bacteria

Co-inertia analysis (CoIA) revealed a significant relationship between AMF genera and fungal classes (RV = 0.29, P = 0.004) (Figure 3.4). The first and second axes representing 80.46 % and 16.80 % of the total projected inertia. A nearly significant relationship was also observed between AMF genera and bacterial classes (RV = 0.20, P = 0.12) (Figure 3.5), with the first and second axes of the CoIA explaining 50.50 % and 44.23 % of the total projected inertia. Figure 3.4 A and Figure 3.5 A show the position of sites on co-inertia axes, where we observe a clear separation between HC and LC sites, and arrow lengths in the LC sites generally being shorter than in HC sites. This indicates that the relationship between community composition of AMF and their associated microorganisms varied across PHP concentrations, which is in line with our PCoA analysis, and that both AMF-fungi and AMF-bacteria relationships are more concordant in the low contaminated sites than in highly contaminated sites.

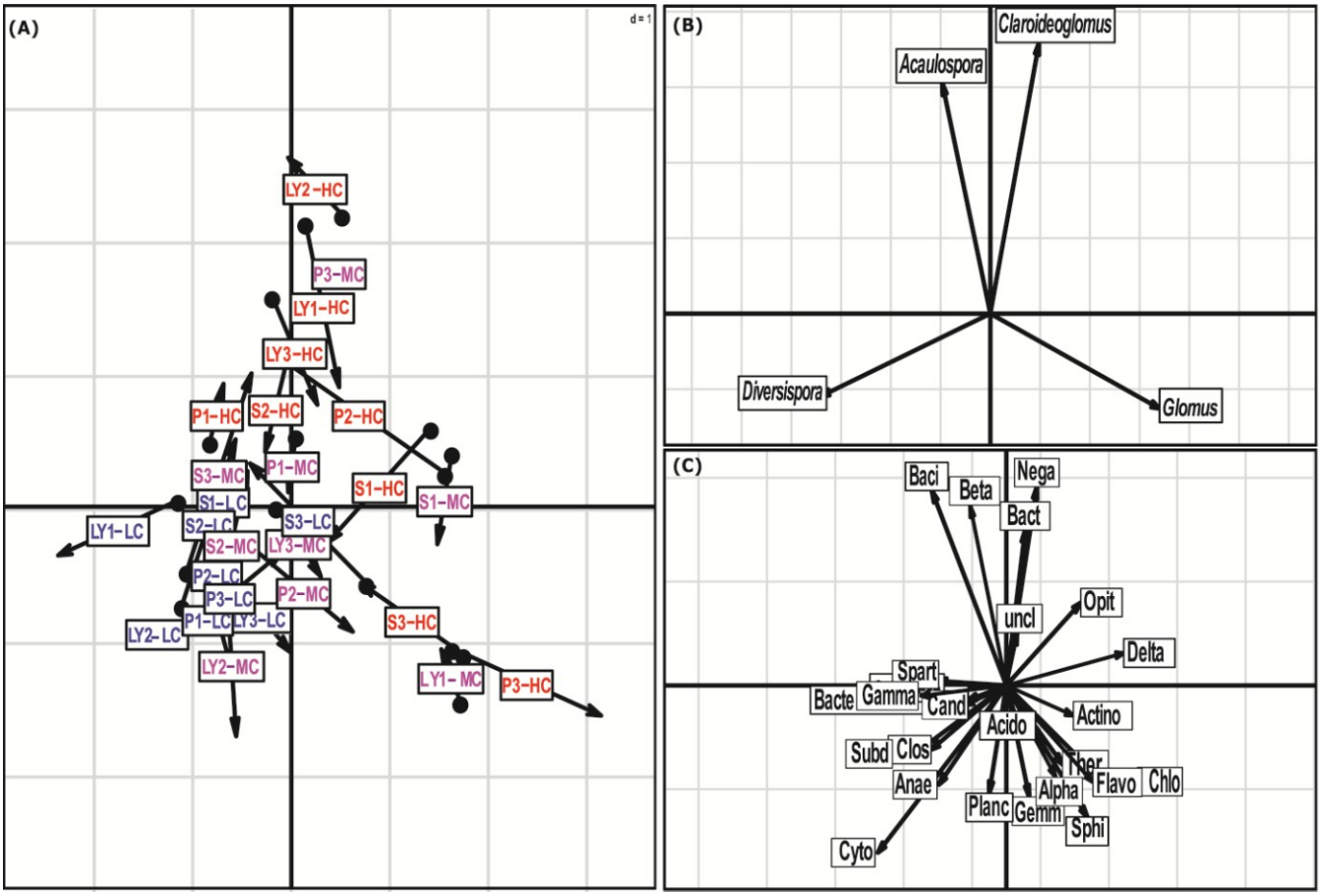


Figure 3.5. Co-inertia analysis showing the relationship between AMF genera obtained from 18S rRNA gene dataset and bacterial classes (RV = 0.20, P = 0.12). A) Projection of both AMF genera and AMF-associated bacterial classes onto the co-inertia plane. B) Projection of AMF genera onto co-inertia plane. C) Projection of AMF-associated bacterial classes onto co-inertia plane.

The base of each arrow (black circle) represents AMF genera, while the tip represents bacterial classes. The lengths of the arrows indicate the degree of concordance between AMF genera and their associated microorganisms. The best concordances were observed in the shortest arrows length. Abbreviations shown in panel C are: Acido: Acidobacteria, Actino: Actinobacteria, Alpha: Alphaproteobacteria, Anae: Anaerolineae, Baci: Bacilli, Bacte: Bacteroidetes_incertae_sedis, Bact: Bacteroidia, Beta: Betaproteobacteria, Cald: Caldilineae (masked by Gammaproteobacteria and Spartobacteria), Cand: Candidatus_Hydrogenedens, Chlo: Chloroflexia, Clos: Clostridia, Cyto: Cytophagia, Delta: Deltaproteobacteria, Flavo: Flavobacteriia, Gamma: Gammaproteobacteria, Gemm: Gemmatimonadetes, Nega: Negativicutes, Opit: Opitutae, Planc: Planctomycetia, Spart: Spartobacteria, Sphi: Sphingobacteriia, Subd: Subdivision3, Ther: Thermomicrobia, uncl: unclassified.

Figure 3.4 B and C shows preferential associations between AMF genera and fungal classes. AMF genera pointing the same direction that AMF-associated fungal classes are positively correlated, whereas those pointing the opposite direction are negatively correlated. These relations are also supported by the Spearman correlations test showing that *Diversispora* and *Glomus* have opposite correlation coefficients across most of the non-AMF fungal classes (Figure S3.7.). *Diversispora* was positively correlated with *Leotiomycetes* (r = 0.38), *Pezizomycetes* (r = 0.38) and *Tremellomycetes* (r = 0.33), while *Glomus* was negatively correlated with these fungal classes. Conversely, *Diversispora* was negatively correlated with *Chytridiomycetes* (r = -0.55), *Dothideomycetes* (r = -0.29) and *Sordariomycetes* (r = -0.37) which were positively correlated with *Glomus* (Figure S3.7.). *P*-values and corrected *P*-values of the different pairwise correlations between AMF genera and fungal classes are shown in the supporting information Table S3.8.

In the case of bacteria, Figure 3.5 B and C shows preferential associations between AMF genera and bacterial classes. Similarly, positive and negative correlations were also observed between AMF genera and some bacterial classes using Spearman correlations test (Figure S3.8.). For example, *Glomus* was correlated with the classes *Betaproteobacteria* (r = -0.37), *Actinobacteria* (r = 0.31), *Bacilli* (r = -0.34) and *Sphingobacteriia* (r = 0.34). *Diversispora* was correlated with *Deltaproteobacteria* (r = -0.33) and *Cytophagia* (r = 0.54). On the other hand, bacterial classes *Gammaproteobacteria*, *Alphaproteobacteria*, *Flavobacteriia* and the *Acidobacteria* groups did not show any significant correlation with any AMF genera. *P*-values and corrected *P*-values of the different pairwise correlations between AMF genera and bacteria classes are shown in the supporting information Table S3.9.

Furthermore, a variance partitioning analysis showed that AMF community structure, PHP concentration and plant species identity explained a total contribution of 21.51 % of the variability in fungal communities (11 % related to plant species identity, 7.6 % related to PHP concentrations and 3.5 % related to AMF) and 24.13 % of the variability in bacterial communities (13.3 % related AMF, 7.4 % related to PHP concentrations and 6.7 % related to plant species identity) (Figure S3.9. and Supplemental Material).

3.6. Discussion

In the rhizosphere, plant roots, arbuscular mycorrhizal fungi (AMF) and the other soil microorganisms share the same microenvironment where they compete for nutrients and exchange complex chemical signaling dialogues, leading to the establishment of saprophytic, parasitic, mutualistic or symbiotic lifestyles. Consequently, the community compositions of AMF and their associated microorganisms are determined in a large part by the different chemical compounds released in the rhizosphere (root exudates and AMF spore exudates). In natural conditions, the community compositions of AMF and their associated microorganisms are also potentially affected by environmental conditions, such as soil composition and climatic conditions. In this study, using the Roche 454 high throughput sequencing platform, we scrutinized the diversity of bacteria and fungi associated with AMF spores collected from the rhizospheric soils of three plant species spontaneously growing in PHP contaminated sites. We found that the assemblages of some AMF and their spore-associated microorganisms were structured across PHP concentration and across plant species. Furthermore, we discovered that the community composition of some AMF-associated microorganisms were intimately linked to the community composition of AMF.

3.6.1. AMF community structure

The taxa belonging to *Glomeraceae* and *Diversisporaceae* families were dominant on the AMF spores identified in both 18S rRNA gene and AMF ITS datasets. *Glomeraceae* were found as the most abundant taxa in the highest contamination level site (HC), while *Diversisporaceae* were the most abundant in the lowest contamination level site (LC). Comparing our results to those of Hassan *et al.* (2014) on rhizospheric soil of 11 *Salix* cultivars collected in a contaminated site nearby the basins of our study, we found that all the AMF families identified in our study were also previously found by these authors, except for *Gigasporaceae*, which were not observed in our study, in both 18S rRNA gene and AMF ITS datasets. Furthermore, in the Hassan *et al.* (2014) study, *Glomeraceae* was also found as the most dominant AMF family in the highly contaminated plots. The presence of *Glomeraceae* in high proportions in the highly PHP contaminated site suggest their ability to tolerate extreme concentrations of PHP. There are several studies reporting that numerous taxa belonging to *Glomeraceae* were tolerant to petroleum hydrocarbon and trace element contaminations, and they suggest that these taxa may accelerate biodegradation of PHP (Vallino *et al.*, 2006; Liu & Dalpé, 2009; Long *et al.*, 2010; Hassan *et al.*, 2011; Aranda *et al.*, 2013; Hassan *et al.*, 2014). For example, Liu and Dalpé (2009) demonstrated that after 12 weeks of growth in a substrate spiked with antracene and phenanthrene, the inoculation of leek roots with *Glomus intraradices* and *Glomus versiforme* isolates reduced concentrations of antracene and phenanthrene by 30 % and 88 %, respectively. Conversely, the presence of *Diversisporaceae* members in higher proportions in the LC site than in the HC site suggest that the species belonging to this AMF family are less tolerant to high concentrations of PHP.

In our study, the effect of plant species on the community structure of AMF was not statistically significant for both ITS and 18S rRNA gene datasets. Previous studies have demonstrated that the host plant species has a substantial effect on the community structures of AMF (Klabi *et al.*, 2014; Toju *et al.*, 2014). This difference can be due to the range of plant species used in each of the different studies or it is possible that the effect of PHP concentration was high enough to mask the smaller effect of host plant on AMF community structures.

We also found that the community composition in the AMF ITS dataset was slightly different from that observed in the AMF 18S rRNA gene dataset. This is not surprising since several studies have demonstrated that the microbial profiles of high throughput sequencing studies varied across the primer pairs and the targeted gene region used (Engelbrektson *et al.*, 2010; Lumini *et al.*, 2010; Gihring *et al.*, 2012). Furthermore, a high genetic polymorphism in ITS regions have been reported within AMF isolates and even within spores of the same isolate (Sanders *et al.*, 1995; Lloyd-Macgilp *et al.*, 1996; Hijri *et al.*, 1999; Redecker *et al.*, 1999).

3.6.2. Associations between AMF spores and microbial communities

We found an high fungal and bacterial diversity associated with AMF spores in PHP contaminated sites. The AMF-associated microorganisms found in this study covered fungal species belonging to *Chytridiomycota* and *Dikarya*, and bacterial species belonging to *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, *Acidobacteria* and *Firmicutes*. The observed AMF-associated microbial communities are in line with what was observed previously by other authors both in *in vitro* conditions and in agricultural soils (Hijri *et al.*, 2002; Mirabal Alonso *et al.*, 2008; Scheublin *et al.*, 2010; Lecomte *et al.*, 2011; Lace *et al.*, 2015; Battini *et*

al., 2016), although much larger bacterial and fungal communities were found here owing to the use of high throughput sequencing. Recently, Iffis *et al.* (2014) identified fungi belonging to *Chytridiomycota*, as well as bacteria belonging mainly to the genera *Sphingomonas*, *Pseudomonas*, *Massilia*, and *Methylobacterium* from vesicles and intraradical spores of AMF collected within *Solidago rugosa* roots growing in the extremely petroleum hydrocarbon-polluted basin on the same site used in our study. However, the roles of the microorganisms associated with AMF and the mechanisms that govern their interactions remain poorly explored particularly in highly stressful conditions such as polluted environments. It has been reported that some genera of *Firmicutes* and *Proteobacteria*, including *Paenibacillus* and *Pseudomonas*, which were identified in this study, are able to enhance AMF colonization by stimulating fungal spore germination, hyphal growth and increasing root branching (Garbaye, 1994; Barea *et al.*, 2005; Frey-Klett *et al.*, 2007; Bonfante & Anca, 2009). Interestingly, some studies have demonstrated that the co-inoculation of AMF with *Proteobacteria* such as *Acinetobacter* sp., *Serratia* sp., and *Sphingomonas* sp. (also identified in the present study) can significantely enhance the rate of petroleum hydrocarbons degradation (Alarcón *et al.*, 2008; Yu *et al.*, 2011; Dong *et al.*, 2014; Xun *et al.*, 2014). In the cases of fungi, Mirabal Alonso *et al.* (2008) isolated two yeast species belonging to *Rhodotorula* and *Cryptococcus* (also identified in the present study) from the spores of *Funneliformis mosseae* (synonym *Glomus mosseae*) with a potential phosphate solubilization activity.

Contrary to the AMF community, which was affected only by the contamination level, the bacterial and fungal communities (excluding AMF) were significantly affected by both plant identity and contamination concentrations (Table 3.1). We found that the class Actinobacteria (represented mainly by the genus of Streptomyces) was present in higher

proportions in spore samples harvested from the MC and HC sites (21.72 % and 6.29 %, respectively) than in LC site (5.12 %). This was expected, since several studies have reported that Streptomyces sp. may grow in PHP contaminated soils and are able to breakdown several recalcitrant petroleum hydrocarbons, such as phenanthrene, pyrene and naphthalene (Balachandran et al., 2012; Bourguignon et al., 2014; Mohamed et al., 2015). Interestingly, Gammaproteobacteria (represented mainly by the genus of *Pseudomonas*) was found in high proportions in both LC (60.69 %) and HC (51.43 %) sites. *Pseudomonas* species are ubiquitous in soils and have a large spectrum of activities in the rhizospheric soils. For example, several studies reported that *Pseudomonas* can potentially degrade a large range of PHP compounds (Kok et al., 1989; Dennis & Zylstra, 2004; Ní Chadhain & Zylstra, 2010), while other studies reported that *Pseudomonas* species are able to establish a symbiotic association with AMF and plant species and may play an important role in improving the AMF colonization, plant growth, nitrogen fixation and phosphate solubilization (Rodríguez & Fraga, 1999; Desnoues et al., 2003; Sharma et al., 2013).

OTUs belonging to classes *Dothideomycetes* and *Chytridiomycetes* were the most dominant AMF-associated fungi in highly and moderately contaminated sites. Previous studies conducted by Bell *et al.* (2014) from contaminated field soils, and by Ferrari *et al.* (2011) from fungal cultures, showed that the abundance of *Dothideomycetes* increased at high hydrocarbon concentrations. However, *Chytridiomycetes* were not reported by Ferrari *et al.* (2011) and Bell *et al.* (2014) studies. The infection of AMF by *Chytridiomycetes* was reported before (Ross & Ruttencutter, 1977; Schenck & Nicolson, 1977; Sylvia & Schenck, 1983; Paulitz & Menge, 1984), though the nature of the relationship between AMF and *Chytridiomycetes* remains unclear. Paulitz and Menge (1984) reported that *Chytridiomycetes* were mainly infecting non-

germinated or dead spores of *Gigaspora margarita*, while Ross and Ruttencutter (1977) and Sylvia and Schenck (1983) found that *Chytridiomycetes* negatively affected the AMF symbiosis functioning.

The effect of plant identity on AMF-associated microorganisms observed here could be determined by the nature of root exudates. In rhizospheric soils, plant roots release a cocktail of complex and specific exudates composed mainly by carbohydrates, amino acids and organic acids, as well as of several signaling compounds by which plants interact with the soil microbes and form a complex rhizospheric microbial community (Bais *et al.*, 2006; Turner *et al.*, 2013a; Rohrbacher & St-Arnaud, 2016). In addition, root exudates are specific for each plant species and they are influenced by several factors such as soil composition, plant age, plant health and climatic conditions, and consequently the microbial communities may change significantly depending on the variation in root exudates composition (Bais *et al.*, 2006; Turner *et al.*, 2013a).

However, the positive and the negative correlations measured between specific AMF taxa and specific AMF-associated microorganisms led us to speculate that AMF identity may also play a key role in shaping the microbial community surrounding their spores. The results obtained from the variance partitioning analysis support this speculation and showed that AMF spores contribute significantly to changes in bacterial community (13.3 %, $P = 0.021$). Whereas, AMF spores were intercorrelated with PHP concentrations and both explained 7 % of the total variation of fungal community (Figure S3.9.).

As plant roots, AMF spores and hyphae may also release exudates and chitin fragments from the out layers of cell walls by which AMF might attract specific microbial communities

(Roesti *et al.*, 2005; Bharadwaj *et al.*, 2011). This could be a strategy by which AMF recruit beneficial microbes, although this hypothesis will need to be tested. Roesti *et al.* (2005) performed four combinations of inoculations between two AMF species (*Glomus geosporum* and *Glomus constrictum*) and two host plant species (*Plantago lanceolata* and *Hieracium pilosella*), and they showed that the AMF associated bacterial communities were more dependent on the AMF spore identity than plant species identity. Recently, Agnolucci *et al.* (2015) found diverse bacterial communities in close association with six different AMF species cultured with the same host plant, under the same environmental conditions and within the same soil. They found the genus *Streptomyces* as the most frequent *Actinobacteria* identified in association with the spores of *Rhizophagus intraradices* (formerly known as *Glomus intraradices*). This was in line with our result that showed positive, yet non-significant, correlation between *Glomus* and *Actinobacteria* ($r = 0.31$, $P = 0{,}12$).

Because AMF show a high genetic and phenotypic polymorphism between species and even between isolates of the same species (Hijri *et al.*, 1999; Redecker *et al.*, 1999; Hijri & Sanders, 2005; Croll *et al.*, 2008; Angelard *et al.*, 2010), the AMF exudates composition may also vary across AMF species or isolates. Consequently, we postulate that AMF taxa recruit different AMF-associated microbes through variation in the composition of their exudates. To date, the most well-known association between specific AMF species and specific microbes is between *Gigaspora margarita* and an the obligatory endobacterium *Candidatus Glomeribacter gigasporarum* (Bianciotto *et al.*, 2003; Ghignone *et al.*, 2012; Salvioli *et al.*, 2016; Vannini *et al.*, 2016), though little is known about the specific AMF-microorganisms interactions and further investigations on this topic are needed.

Overall, we demonstrated that AMF may undergo complex interaction not only with plants, but also with soil microbes, and therefore we assumed that AMF can be seen as corner stone interface bridging and extending the complex root microbial assemblages beyond roots and their rhizospheres.

3.7. Acknowledgements

This work was supported by the Genome Quebec, Genome Canada and The *Natural Sciences and Engineering Research Council* of Canada (NSERC) which are greatly acknowledged. We thank Stéphane Daigle for assistance in statistics, Rim Klabi for help in sequence analyses, Louise Pelletier for her assistance in scanning electron microscopy. We thank Terrence Bell and Karen Fisher-Favret for commenting on and editing the manuscript. We also thank three anonymous reviewers for their helpful comments.

Mise en contexte chapitre 4 - Comparaison entre les communautés microbiennes associées aux CMA et celles du sol et des racines de plantes.

Dans le chapitre précédent, le séquençage à haut débit des régions 18S d'ARNr des CMA, 16S d'ARNr bactériens et des ITS fongiques nous a permis de faire une étude globale sur la biodiversité des CMA et des microorganismes qui sont leurs associées dans un site hautement contaminé par les HP. Nos résultats nous ont permis de constater que les communautés microbiennes associées aux CMA sont influencées par les niveaux de contamination et les espèces de plantes hôtes. De plus, les corrélations observées entre quelques espèces de CMA et de microorganismes nous laissent suggérer que les CMA peuvent aussi jouer un rôle clé dans la structure des communautés de ces microorganismes.

Dans le présent chapitre, nous avons conservé la même démarche expérimentale que dans le chapitre 3, mais cette fois-ci nous nous sommes focalisés sur l'étude de la biodiversité des bactéries et des champignons présents dans les sols et les racines afin de les comparer avec celle retrouvée en association avec les CMA.

Les résultats de cet article sont en préparation pour la publication dans une revue scientifique internationale.

Ma contribution dans cet article est la participation à la mise en place de l'expérience et la réalisation de la totalité des travaux d'échantillonnage et de laboratoire. J'ai également fait toutes les analyses bioinformatiques et statistiques. J'ai aussi participé à la rédaction et j'ai fait toutes les illustrations présentées dans l'article.

Chapitre 4 - Petroleum hydrocarbon concentrations and plant species identity influence bacterial and fungal community structures across soil and roots

Bachir Iffis, Marc St-Arnaud and Mohamed Hijri

Institut de Recherche en Biologie Végétale, Département de Sciences Biologiques, Université de Montréal, 4101 Sherbrooke est, Montréal (Québec) H1X 2B2, Canada

En préparation pour publication

Author contributions : Conceived and designed the experiments: BI MH. Performed the experiments: BI. Analyzed the data: BI. Contributed reagents/materials/analysis tools: BI. MH. Wrote the paper: BI MH MSA.

4.1. Abstract

Phytoremediation is a promising *in situ* green technology based on the use of plants to cleanup soils from organic and inorganic pollutants. Microbes, particularly bacteria and fungi, that closely interact with plant roots play key roles in phytoremediation processes. In polluted soils, the root-associated microbes contribute to alleviation of plant stress, improve nutrient uptake and may either degrade or sequester a large range of soil pollutants. Therefore, improving the efficiency of phytoremediation requires a thorough knowledge of the microbial diversity living in close association with plant roots in both the rhizosphere and the endosphere. This study aims to assess fungal ITS and bacterial 16S rRNA gene diversity using high-throughput sequencing in rhizospheric soils and roots of three plant species (*Solidago canadensis, Populus balsamifera* and *Lycopus europaeus*) growing spontaneously in three petroleum hydrocarbon polluted sedimentation basins. Because we have previously conducted a thorough study on the diversity of bacteria and fungi associated with arbuscular mycorrhizal fungi (AMF) harvested from the same areas, the microbial community structures of rhizospheric soils and roots were compared with those of microbes associated with AMF spores in order to verify whether AMF are able to recruit specific microbial communities on the surface of their spores and mycelia. Our results showed a difference in OTU richness and community structure composition between soils and roots for both bacteria and fungi. We found that petroleum hydrocarbon pollutant (PHP) concentrations have a significant effect on fungal and bacterial community structures in both soils and roots, whereas plant species identity showed a significant effect only on the roots for bacteria and fungi. Our results also showed that the community composition of bacteria and fungi in soil and roots varied from those associated with AMF spores harvested from the same plants. This let us to speculate that

in PH contaminated soils, AMF release chemical compounds by which they recruit beneficial microbes in order to tolerate or degrade the PH pollutants present in the soil.

4.2. Keywords

Petroleum hydrocarbon pollutants; 454 high throughput sequencing; phytoremediation; AMF-associated microbes.

4.3. Introduction

During the past century, industrial production, urbanization, energy consumption, transportation and human population have expanded exponentially, resulting in increased soil, water and air pollution, which in turn has placed the environment under substantial pressure (Samanta *et al.*, 2002; Chen & Kan, 2008). Together, these factors produced a large number of highly polluted sites all over the planet, containing usually complex mixtures of toxic and carcinogenic, organic and inorganic compounds. Organic contaminants such as PAHs and PCBs are known mutagens and carcinogens that enter the food chain together with lipophilic compounds (Boffetta *et al.*, 1997; Henner *et al.*, 1997; Poirier, 2004). Inorganic contaminants mainly consist of metalloids and trace metals with soil retention times of up to thousands of years. Like organic compounds, they reduce plant growth, negatively impact the soil microbiota (McGrath *et al.*, 2001), and decrease the quality of the environment (Gremion *et al.*, 2004) to such an extent that they pose serious health risks to humans and animals.

Polluted sites may be cleaned by physico-chemical strategies including excavation and storage, washing, and chemical treatment. Yet, most *ex situ* treatments only contain contamination without eliminating it. They damage or even destroy soil microbial communities, and are unfit for application over large areas because they are prohibitively

expensive. An alternative, most promising *in situ* approach for multi-contaminated sites is phytoremediation (Salt *et al.*, 1995; Salt *et al.*, 1998; Peuke & Rennenberg, 2005), which uses plants and their associated soil microbial communities (fungi and bacteria) to accumulate pollutants within the plants, and/or degrade them in the soil (Peuke & Rennenberg, 2005; Pilon-Smits, 2005; Bieby Voijant *et al.*, 2011).

In the past few years, phytoremediation method has become increasingly popular owing to its efficiency, cost effectiveness and respect for the integrity of the soil structure and biology. However, in many cases, phytoremediation is best achieved through a complex interaction between plants and the myriad of bacteria and fungi living in the rhizospheric soil and inside plant roots (endophytic microorganisms) (Pilon-Smits, 2005; Jing *et al.*, 2007; Ma *et al.*, 2016; Rohrbacher & St-Arnaud, 2016; Thijs *et al.*, 2016). Therefore, improving the efficiency of phytoremediation requires a thorough knowledge of the microbial diversity living in association with plant roots (either in the rhizosphere or the endosphere), and of their interactions with plants. Several studies demonstrated that, in contaminated soils, the root exudates released in the plant rhizosphere promote the selection of microorganisms able to degrade pollutants and stimulate the expression of several genes involved in xenobiotic compound degradation (Bell *et al.*, 2014; Yergeau *et al.*, 2014; Pagé *et al.*, 2015; Rohrbacher & St-Arnaud, 2016; Thijs *et al.*, 2016). For example, Pagé *et al.* (2015) used a metatranscriptomic approach to compare the gene expression of ten oxygenases related to petroleum hydrocarbon degradation between the bulk and rhizospheric soils of *Salix purpurea.* They found that among the ten genes examined, four of them were significantly over-expressed in rhizospheric soil compared with bulk soil. Bell *et al.* (2014) studied the fungal and bacterial diversity from the rhizospheric soils of eleven cultivars of willow planted in

three hydrocarbon contaminated sites, and they found that the abundance of some petroleum hydrocarbon-degrading microorganisms, such as some classes of *Proteobacteria* (*Alpha*, *Beta* and *Gamma- proteobacteria*) and *Dothideomycetes* (fungi), was significantly enhanced in the highly contaminated plots compared with the low and non-contaminated plots.

Arbuscular mycorrhizal fungi (AMF) are well known as soil fungi able to establish a mutualistic symbiosis with over than 80 % of the land plants (Smith & Read, 2008). In exchange for carbon resources, AMF provide the host plants with nutrients and protect them against soil-borne pathogens (St-Arnaud & Vujanovic, 2007; Smith & Read, 2008; Ismail *et al.*, 2011). In addition, many reports have shown that AMF may play an important role in soil phytoremediation processes (Liu & Dalpé, 2009; Wu *et al.*, 2009; Gao *et al.*, 2011b; Hassan *et al.*, 2013). In the soil surrounding plant roots, AMF share the same micro-environment with other rhizospheric microorganisms and several studies have shown that AMF species collaborate with some of these microorganisms in phytoremediation process (Alarcón *et al.*, 2008; Liu & Dalpé, 2009; Teng *et al.*, 2010). However, AMF also harbor their own hyposphere microorganisms on the surface of their spores and mycelium (Hijri *et al.*, 2002; Bonfante & Anca, 2009; Scheublin *et al.*, 2010; Lecomte *et al.*, 2011), and the role of these microbes in phytoremediation processes is unknown.

Using 454 sequencing, Iffis *et al.* (2016) have conducted a study of bacteria and fungi associated to AMF spores harvested from soil collected from the rooting zone of three plant species (*Solidago canadensis, Populus balsamifera* and *Lycopus europaeus*) growing spontaneously in waste decantation basins of a former petrochemical plant. These authors have found a large diversity of bacteria and fungi in association with the AMF spores. They also found that the AMF-associated fungal and bacterial communities were significantly affected

by both PHP concentrations and plant species identity. Furthermore, Iffis *et al.* (2016) observed that some AMF taxa were either positively or negatively correlated with some fungal and bacterial groups, suggesting that AMF may also play a role in shaping the microbial communities associated with their spores. Similarly, Iffis *et al.* (2014) showed that the intraradical propagules (vesicules and spores inside plant roots) of AMF extracted from *Solidago rugosa* roots, sampled in a PHP polluted site, also harbored a large diversity of bacteria and fungi.

Based on these studies, we inferred that plant species, AMF community and PHP concentrations are among the major driving forces that shape microbial communities living in the rhizosphere and endosphere of plants growing in PHP polluted sites. The current study aims to understand the contribution of each of these factors on bacterial and fungal community structures in rhizospheric soils and roots sampled in polluted sedimentation basins. Our objectives were: (i) to assess the bacterial and fungal diversity associated with roots and their rhizospheric soil from mycorrhizal plants spontaneously growing in waste decantation basins of a former petrochemical plant; (ii) to test the effects of petroleum hydrocarbon pollutants (PHP) and plant species identity on the microbial community structure in the soil and plant roots; (iii) and to compare the microbial community structure of soil and roots with the microbial community structure associated with the AMF spores in order to verify the hypothesis that AMF are able to recruite specific microbial communities on the surface of their spores and mycelia.

To do so, soil and root samples used in Iffis *et al.* (2016) were subjected to DNA extraction, PCR amplifications targeting the ITS regions for fungi and 16S rRNA gene for

bacteria, then the PCR products were sequenced using the 454 FLX+ high throughput sequencing platform to profile the microbial communities structure.

Overall, our results show a difference in OTUs richness and community structure composition between soil and roots for both bacteria and fungi. We also found that PHP concentration have a significant effect on the fungal and bacterial community structures in both soil and roots while, plant species identity had a significant effect only on the root bacteria and fungi. Furthermore, the comparison between the results of this study and Iffis *et al.* (2016) study showed that the microbial community structures found in soil and root differed from those found in association with the AMF spores harvested from the same samples. These results support the hypothesis that AMF are able to recruit specific microbial communities on the surface of their spores and mycelium. According to our knowledge, this is the first research work devoted to compare between the microbial communities of soil, roots and in association with AMF spores in PH contaminated sites.

4.4. Materials and methods

4.4.1. Experimental design and sampling

Field sampling and experimental design were previously described in Iffis *et al.* (2016). Briefly, *Solidago canadensis, Populus balsamifera* and *Lycopus europaeus* are three plant species naturally growing in three petroleum contaminated basins of an former petrochemical plant located on the south shore of the St-Lawrence River near Montreal, Quebec, Canada (45.70 N, 73.43 W) (Desjardins *et al.*, 2014). The root system of three individual plants with their surrounding soils were collected for each plant species and from each basin, totalling 27 samples (three basins × three plant species × three replicates).

The root system from each plant was washed several times in tap water until elimination of all soils debris attached, and then cut into 1 cm long pieces. Since a part of the root samples was already used for microscopic examination, the remaining amounts were insufficient for individual DNA extractions. Therefore, the root replicates of each plant species were pooled in 10 ml tubes, and ground in liquid nitrogen with a mortar and pestle. Then, 500 mg aliquots of the ground root material were collected in 1.5 ml tubes, totaling 9 tubes for plant roots. As the remaining soil amount was sufficient for individual DNA extraction, the soil samples were kept separate for each individual per plant species from each basin, so there were a total of 27 soil samples. For each replicate, the soil surrounding the roots was collected in plastic bag, homogenized by shaking, then 500 mg aliquots of this soil were collected in 1.5 ml tubes for DNA extraction. Root and soil tubes were conserved at -20 °C until use.

The analysis of polycyclic aromatic hydrocarbon (PAH) and total alkane (C10–C50) concentrations from the soil of the three contaminated basins was carried out using a commercial laboratory service (Maxxam Analytical Laboratories, Montreal, QC), and the results are available in Table S3.10. of Iffis *et al.* (2016). Basin 1 was the most highly contaminated with mean concentration of alkanes (C10–C50) equal to 2200 mg/kg of soil (termed as HC), basin 2 was the least contaminated with mean concentration of alkanes (C10–C50) equal to 153 mg/kg of soil (termed as LC), while the basin 3 was moderately contaminated with mean concentration of alkanes (C10–C50) equal to 800 mg/kg of soil (termed as MC) (Iffis *et al.*, 2016).

4.4.2. DNA extraction

DNA extraction was performed in both soil and roots samples using the NucleoSpin Soil Kit (MACHEREY-NAGEL, USA) following the manufacturer's instructions. DNA was eluted in a 50 µl of elution buffer, and was stored at -20 °C until use.

4.4.3. Polymerase chain reactions

To identify the bacterial taxa, soil and root DNA samples were subjected to PCR amplification targeting a partial 16S rRNA gene fragment with primer pair UnivBactF 9 and BSR534/18 (Bell *et al.*, 2014). To identify the fungal taxa, the ITS regions were targeted using primer pair ITS1F and ITS4 (Bell *et al.*, 2014). For the soil DNA samples, PCR amplifications were performed once on each DNA sample, totaling 27 PCR tubes of soil samples. For the root DNA, since plant root replicates were pooled, PCR amplifications were performed in triplicate on each DNA sample, also totaling 27 PCRs tubes.

Both 16S and ITS primers were tagged with adaptors and unique multiplex identifier (MID) tags from the extended MID set recommended by Roche Diagnostics (Roche, 2009). PCRs were performed in 50 µl volumes containing 5 µl of 10× PCR buffer, 0.2 mM of dNTP mix, 1 µl of BSA (100 mg/ml), 1 µl $MgCl_2$ (25mM), 0.4 mM of each primer, 2 µl of DNA template and 1 U of *Taq* DNA polymerase (QIAGEN, Toronto, ON, Canada). PCRs were run on a thermal cycler Pro S thermocycler (Eppendorf, Mississauga, ON, Canada) using an initial denaturation at 95 °C for 5 min followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min, and a final elongation step at 72 °C for 10 min. After electrophoresis separation and UV light visualisation, the PCRs products were purified with the QIAquick Gel Extraction Kit (QIAGEN, Toronto, ON, Canada) following the manufacturer's instructions. Then, DNA concentrations of the purified products were measured using the Qubit 2.0 fluorometer (Life

Technologies, Burlington, ON, Canada). Four pools were prepared by mixing equimolar volumes of PCR products. The four pools contained 16S rRNA gene amplicons of soil samples, 16S rDNA amplicons of root samples, ITS amplicons of soil samples and ITS amplicons of root samples. These pools were sent for sequencing to the Genome Quebec Innovation Centre using the Roche 454 FLX+ pyrosequencing platform (Roche, Branford, CT, USA). One eighth of sequencing plate was used for sequencing each pool.

4.4.4. Bioinformatic processing

Sequence processing was performed in Mothur (v.1.34.4) (Schloss *et al.*, 2009) as described in Mothur wiki (http://www.mothur.org/) and (Bell *et al.*, 2015) with some minor modifications. Briefly, for both soil and root 16S rRNA gene sequences, the ".sff" files of the different samples were first merged in one ".sff" file, then ".qual" and ".fasta" files were obtained from the ".sff" file using "merge.files" and "sffinfo" commands, respectively. Low quality and short sequences were removed using the "trim.seqs" command, with the following parameters: maxambig=0, maxhomop=8, bdiffs=1, pdiffs=2, qwindowaverage=30, qwindowsize=50, minlength=300 and then, we reduced the dataset to only unique sequences using the "unique.seqs" command. After sequence alignments against the Mothur interpreted Silva bacterial database using "align.seqs", the non-aligned sequences were removed using the "screen.seqs" command with the following criteria: start=1044, optimize=end, criteria=95. Here, we used reference-based clustering instead of *de novo* clustering because, it is difficult to process a large amount of sequences (e.g. sequences produced by next generation sequencig) with *de novo* clustering, owing to their intensive computational algorithms. Prior to sequence classifications, the datasets were first subjected to a second simplification using the "unique.seq" command, followed by the commands "pre.cluster" and "chimera.uchime" to

reduce the sequencing errors. Sequence classifications were carried out with the Mothur implementation of the RDP database using the "classify.seqs" command. Sequences identified as "Mitochondria", "Chloroplast", "Archaea", "Eukaryota", or "unknown" were removed using the "remove.lineage" command. The sequences matched with Archaea were removed because the primer sets used in this study are universal primers for eubacteria and therefore, did not allow to recover all the diversity of Archaea. Distance matrices were generated with "dist.seqs" command and OTUs were obtained using the "cluster.split" command (method=average, processors=2, splitmethod=classify, large=T). Removal of the singleton reads was carried out using the "split.abund" command and then, to have an equal number of reads per sample, datasets were standardized by a random sub-sampling using the "sub.sample" command. OTU tables at 97% similarity were generated following the steps for "create.database".

Overall, the steps of soil and root ITS datasets processing were similar to those of 16S rRNA gene dataset processing. Most of the modifications were introduced owing to the absence of reference database for ITS sequence alignments. After elimination of low quality and short sequences by the "trim.seqs" command (maxambig=0, maxhomop=8, bdiffs=1, pdiffs=2, qwindowaverage=30, qwindowsize=50, minlength=250), we ran "chop.seqs" and "chimera.uchime" commands to standardize sequence length (numbases=249) and reduce the sequencing errors, respectively. Then, the ".fasta" files were clustered into OTUs at 97% identity using the CD-HIT software (Li & Godzik, 2006) and reformatted to ".list" file in order to continue Mothur processing. Since Iffis *et al.* (2016) used CD-HIT software for clustering, we used the same software in this study in order to keep the same parameters for comparisons.The remaining steps of the ITS datasets processing was similar to 16S rRNA

gene dataset processing, except sequence classification was performed using the UNITE reference database, and the "remove.lineage" command was not used in ITS datasets processing.

4.4.5. Statistical analysis

All statistical analyses were performed in R (version 3.1.1). To compare the diversity and OTU richness between soil and root datasets, Student's t-tests or Wilcoxon tests were carried out on Shannon diversity indices and the Chao richness estimator using the "Rcmdr" package. Student's *t*-tests were performed on the fungal Chao values (normally distributed), while Wilcoxon tests were performed on the bacterial Chao and Shannon values, and on the fungal Shannon values (non-normally distributed residuals). Student's t-tests or Wilcoxon tests were also performed in order to compare the fungal and bacterial relative abundances between soil and roots. To verify the efficiency of our sampling efforts and sequencing depth, rarefaction curves were drawn for each individual sample using the "rarefy" function in the "vegan" package. Depending on the normal distribution of the residuals, the effect of contamination levels and plant species identity on Shannon diversity indices of fungi and bacteria were tested by ANOVA or Kruskal-Wallis tests using the "Rcmdr" package in R (ANOVA tests were carried out on soil fungi, root fungi and soil bacteria, while Kruskal-Wallis tests were performed on root bacteria.

To test the effect of contamination levels, plant species and biotopes (soil roots and AMF spores) on the bacterial and fungal community structures, PERMANOVA analysis were performed using the "adonis" function in the "vegan" package on Bray–Curtis values obtained from the community structure matrices previously normalized by the Hellinger transformation. The Hellinger transformations were carried out in order to alleviate the effect the highest

abundances and the double zeros problemes (Legendre & Legendre, 2012). For the PERMANOVA carried out across biotopes, since the microbial datasets of soil, roots and AMF spores were obtained from different sequencing pools with different sequencing depth, their abundances were transformed on percentages and the taxonomic data were summed at the genera level, prior to the merging of matrices and PERMANOVA analyses. For the PERMANOVA performed to compare the fungal communities of soil and roots with those of AMF spores, all the taxa matching with Glomeromycota were removed from soil and root matrices, prior to the analysis. To test the homogeneity of dispersion of the different communities against PHP concentration and plant species identity, beta-dispersion analyses were performed on the Bray–Curtis matrices using the "betadisper" function in the "vegan" package. In order to reveal which fungal or bacterial taxa were affected by the contamination levels or plant species, Kruskal–Wallis tests or ANOVA were performed on the abundances at class level and on the most abundant thirty OTUs of fungi and bacteria in both soil and root datasets. Non-metric multidimensional scalings (NMDS) were performed in order to visualize the effects of contamination levels, plant species and biotopes on community composition. NMDS ordinations were calculated from the Bray–Curtis matrices using the "metaMDS" function from the vegan package. Relative abundances were calculated with Excel software on the different datasets in order to visualize the percentages of taxonomic affiliations across biotopes, contamination levels and plant species. Krona charts were calculated using the KronaTools available from (https://github.com) (Ondov *et al.*, 2011) in order to compare the AMF-associated microbial communities (Iffis *et al.*, 2016) with the soil and roots microbial communities.

4.5. Results

4.5.1. Soil microbial diversity versus root microbial diversity

After quality filtering and standardizing the number of sequences in the different datasets, the soil 16S rRNA gene dataset allowed us to retrieve a total of 23085 reads (855 per sample) which were assigned to 4083 OTUs, while a total of 25839 reads (957 per sample) were obtained from the soil ITS dataset which were assigned to 215 OTUs. For the root 16S rRNA gene dataset, we retrieved a total of 2403 reads (89 per samples) which were assigned to 820 OTUs, while a total of 13716 reads (508 per samples) were obtained from the root ITS dataset and assigned to 188 OTUs. Since the different datasets were standardized by subsampling to the sample size that showed the lowest number of reads, the number of reads retrieved in each dataset was relatively low in comparison with what 454 sequencing technology can produce. Furthermore, the presence of PCR inhibitors in our samples may be one of the major factor that affected the sequencing depth.

Rarefaction curves showed that the sampling effort for fungi was close to saturation for the all samples with Good's coverage values ranged between 0.97 and 0.99 for soil fungi, and 0.94 and 0.98 for root fungi. Contrary to fungi coverage, the sampling efforts were relatively low yielded for bacteria, in particular the samples related to *L. europaeus* and *P. balsamifera*. The Good's coverage values of bacteria ranged between: 0.68 and 0.81 for soil bacteria, 0.22 and 0.70 for root bacteria, (Figure 4.1. and Table S4.1.).

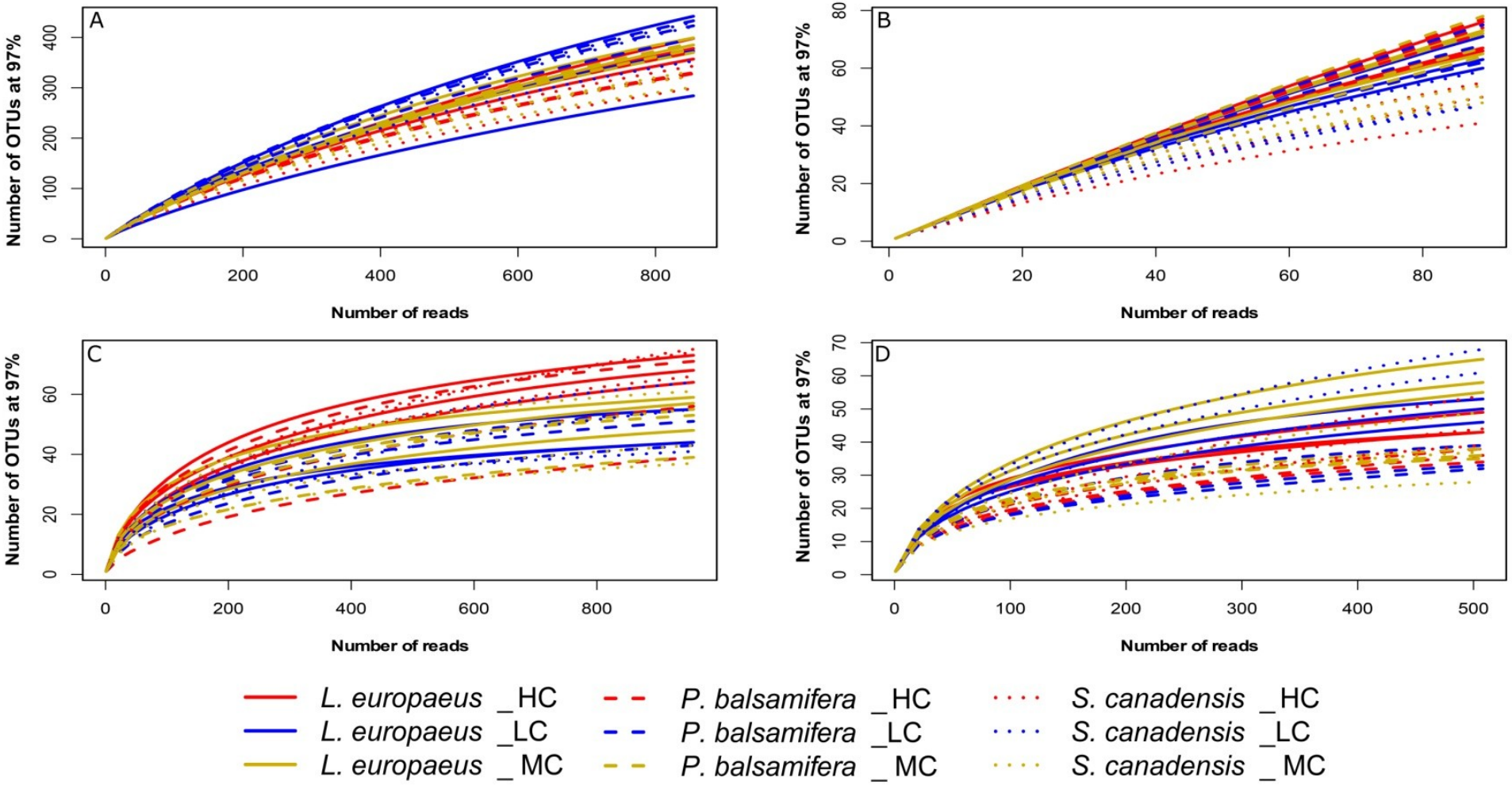


Figure 4.1. Rarefaction curves of OTUs for individual samples across the different datasets: (A) soil bacteria, (B) root bacteria, (C) soil fungi, (D) root fungi.

The bacterial OTU diversity and richness of bacteria were significantly higher in soils than in roots (Wilcoxon test, $P < 0.0001$) for both Shannon diversity indices and Chao richness estimators (Figure 4.2. A and B). For fungi, Wilcoxon test on Shannon diversity indices showed no significant difference in fungal diversity between soil and root datasets ($P = 0.15$). However, a significant difference in OTU richness was observed between the soil and root samples for fungi (Student's *t* test, $P = 0.05$), with Chao values higher in the soil samples (mean Chao value = 68.66) compared to root samples (mean Chao value = 59.58) (Figure 4.2. A and B).

PERMANOVA analyses showed that the bacterial and fungi community structures present in soil were significantly different from the bacterial and fungi communities found in association with plant roots ($P < 0.001$). For both bacteria and fungi, NMDS ordinations showed a clear change in community structures across soil and roots (stress value = 0.14 and 0.22, respectively) (Figure 4.3. A and Figure 4.4. A)

BLAST searches of the bacterial 16S rRNA gene sequences showed that, at class or phylum level, the bacteria profiles were similar in soil and root datasets. The most abundant bacterial taxa identified in the two datasets belong to classes *Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Actinobacteria,* and phylum *Acidobacteria*. However, except for *Alphaproteobacteria* and *Betaproteobacteria*, which were found approximately in similar proportions in soil and root datasets (36 % and 33 % for *Alphaproteobacteria,* and 14 % and 12 % for *Betaproteobacteria*), the proportions of the other most abundant bacteria groups were different between the soil and root datasets. *Acidobacteria* and *Deltaproteobacteria* had higher proportions in the soil dataset (11 % and 5.25 %, respectively) compared to the root dataset (4 % and 3.41 %, respectively) (Student's *t* test, $P <$

0.05) while, the percentages of *Actinobacteria* (represented mainly by *Streptomyces*) and *Gammaproteobacteria* were higher in the root dataset (19 % and 14 %, respectively) than in the soil dataset (6% and 9%, respectively) (Wilcoxon test, $P < 0.0001$ for *Actinobacteria* and Student's *t* test = 0.15 for *Gammaproteobacteria*) (Figure S4.1. A and B). Moreover, within each bacterial group, most OTUs found in the root dataset were different from those found in the soil dataset (Table S4.2. and, Figure S4.2. A and B). For instance, in the root dataset, the OTUs related to *Alphaproteobacteria* were represented mainly by *Bradyrhizobium*, *Skermanella*, *Sphingobium* and *Hoeflea*, while in the soil dataset, *Alphaproteobacteria* were dominated by *Sphingomonas*, *Skermanella* and *Dongia*. Similarly, *Betaproteobacteria* was represented mainly by *Ideonella*, *Duganella* and *Limnobacter* in the root dataset, while it was dominated by *Caenimonas*, *Burkholderiales* and *Ferrovum* in the soil dataset.

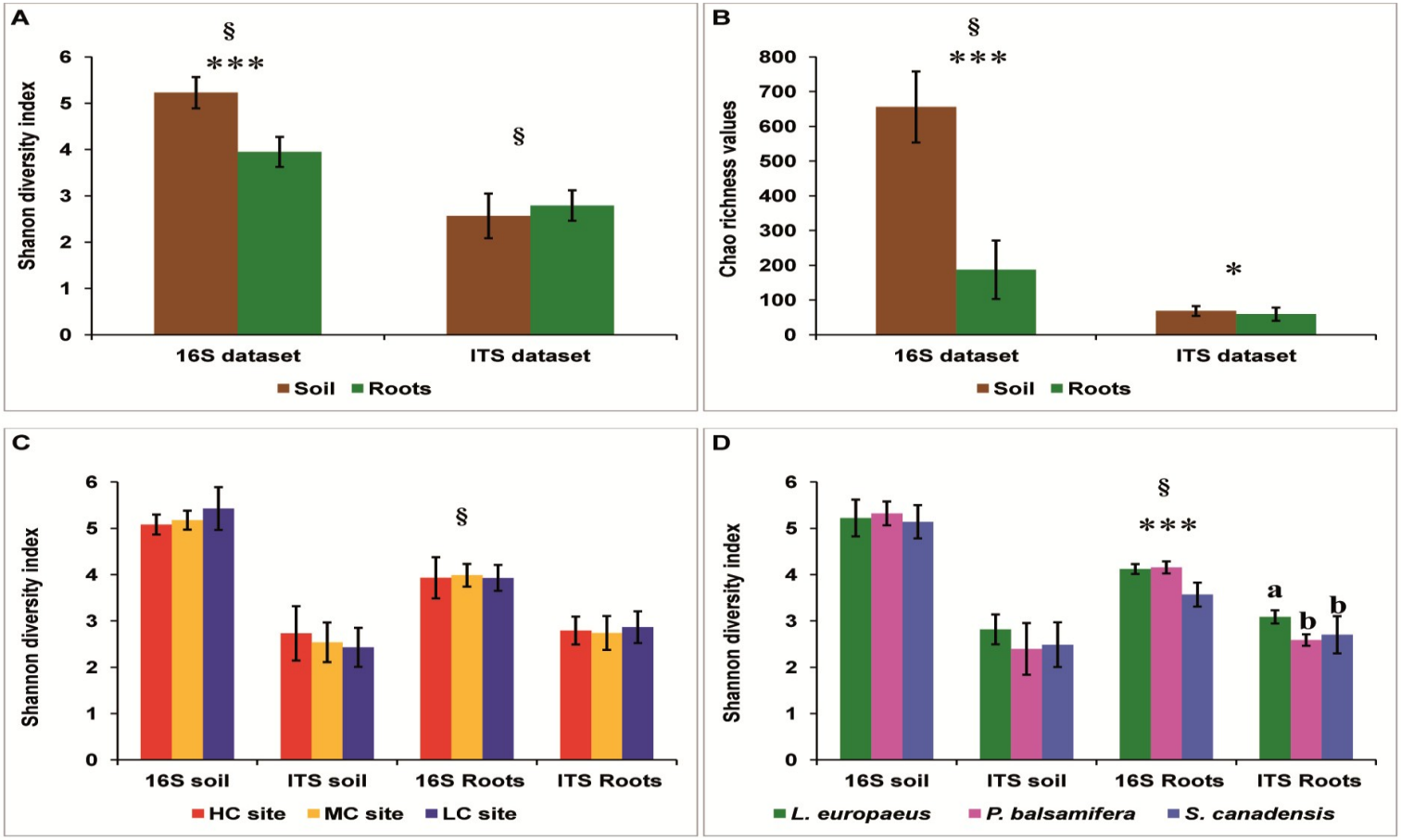


Figure 4.2. (A) Comparison of Shannon diversity indices of bacteria and fungi across soil and roots. (B) Comparison of Chao estimator values of bacteria and fungi across soil and roots. (C) Comparison of Shannon diversity indices of the different pyrosequencing datasets across contamination concentrations. (D) Comparison of Shannon diversity indices of the different pyrosequencing datasets across plant species.

Acronyms HC, CM and LC mean high contamination, moderate contamination, and low contamination, resperctively. (∗): significant at 5 %; (∗∗∗): significant at 0.1 %; (§): *P* values calculated using Wilcoxon or Kruskal-Wallis tests.

For the ITS sequences, BLAST searches showed that almost all fungal taxa identified in the soil dataset were also identified in the root dataset. The major part of the fungal taxa found in the two datasets belong to the classes *Dothideomycetes*, *Sordariomycetes*, *Agaricomycetes*, *Chytridiomycetes* and *Glomeromycetes*. *Dothideomycetes*, *Agaricomycetes* and *Chytridiomycetes* were found in similar proportions in the root and soil datasets. *Dothideomycetes* represented 23 % and 24 % of the OTUs in soil and root datasets, respectively. *Agaricomycetes* represented 10 % and 7 % of OTUs in the soil and root datasets, respectively. Similarly, *Chytridiomycetes* represented 5 % and 4 % of OTUs in the root and soil datasets, respectively. However, the proportions of *Sordariomycetes* and *Glomeromycetes* were clearly different between soil and root samples (Wilcoxon test, $P \leq 0.0001$). *Sordariomycetes* were found as the most dominant taxa in the soil dataset, with a percentage of 37 % of OTUs, whereas they represented only 5 % of the root dataset. *Glomeromycetes* represented 20% of the root dataset, but only 1 % of the fungal sequences in the soil dataset (Figure S4.1. C and D).

4.5.2. PHP and plant species identity effects on soil and root bacterial diversity

ANOVA tests revealed a nearly significant effect of the contamination level on Shannon diversity indices of bacteria in the soil dataset ($P = 0.077$), with a higher diversity in the least contaminated (LC) site than in most highly contaminated (HC) and moderately contaminated (MC) sites. On the other hand, there was clearly no effect of contamination on the bacterial diversity indices in the root dataset (Table 4.1 and Figure 4.2. C). However, when comparing the shifts in Shannon diversity indices by plant species, a highly significant effect of plants species identity on bacterial diversity was observed in the root dataset ($P < 0.001$), with a higher diversity of bacteria in *L. europaeus* and *P. balsamifera* than in *S. canadensis*. In

contrast, there was no effect of plant species identity on the bacterial diversity in the soil dataset (Table 4.1 and Figure 4.2. D).

Table 4.1. *P*-values of ANOVA and PERMANOVA analyses calculated on Shannon diversity indices and community structures of the different pyrosequencing datasets.
n = 9 for soil bacteria, root bacteria, soil fungi and root fungi. The Bolded values are significant at $P < 0.05$. (§): *P*-values calculated using Kruskal-Wallis tests.

	Contamination level				Plant species			
	ANOVA on Shannon index	**PERMANOVA on the community structure**	**Beta-dispersion**	**PERMANOVA R squared**	**ANOVA on Shannon index**	**PERMANOVA on the community structure**	**Beta-dispersion**	**PERMANOVA R squared**
Soil bacteria	0.0777	**0.011**	0.25	0.05	0.544	0.428	0.49	0.03
Root bacteria	§0.7242	**0.034**	0.46	0.05	§**0.0001**	**0.001**	**0.031**	0.07
Soil fungi	0.424	**0.002**	0.72	0.08	0.151	0.37	0.10	0.03
Root fungi	0.714	**0.006**	0.51	0.09	**0.000965**	**0.002**	0.06	0.11

PERMANOVA analysis revealed that the community structure of bacteria was significantly affected by the contamination level for both soil and root datasets (P = 0.01 and 0.032, respectively). However, a significant effect of plant species identity on the bacterial community structure was observed only on the root dataset (P = 0.001), while no effect of plant species was observed on bacteria in the soil dataset (P = 0.41) (Table 4.1).

The variations in diversity and community composition of bacteria were corroborated by the non-metric multidimensional scaling (NMDS) plots in both soil and root datasets. In NMDS plots performed across contamination levels, a clear separation of the bacterial communities was observed between the LC and HC sites in both soil and root datasets (stress value = 0.15 and 0.22, respectively) (Figure 4.5. A and B). However, the NMDS plot performed across plant species showed that the bacterial communities were determined by plant species identity only in the root dataset, where bacteria from *S. canadensis* roots grouped apart those from the other plant species (stress value = 0.15 and 0.22, respectively) (Figure 4.6. A and B).

Kruskal-Wallis tests performed on the soil bacteria OTUs showed clearly that the contamination level had a stronger effect on the bacterial composition than plant species (Table S4.2.). Among the most abundant 30 OTUs, 17 were significantly affected by the contamination levels, two by plant species, and two by both contamination levels and plants species, while no effect was found on the remaining nine OTUs. Most of the OTUs affected by the contamination belong to classes *Alphaproteobacteria* (*Sphingomonas*, *Skermanella*, *Dongia*, *Rhizobiales*), *Betaproteobacteria* (*Caenimonas*, *Burkholderiales*, *Ferrovum*, *Comamonadaceae*), *Gammaproteobacteria* (*Xanthomonadales*, *Thermomonas*, *Steroidobacter*) and *Acidobacteria* groups (Table S4.2.).

At the class level, The proportion of *Alphaproteobacteria* and *Acidobacteria* groups were significantely increased in the HC (40.8 % and 11.6 %) and MC (37.7 % and 12.9 %) sites more than in the LC site (29.8 % and 8.7 %) (ANOVA, P = 0.012 and 0.002). *Betaproteobacteria* showed also a high proportion in the HC and MC sites (13.3 % and 15 %) than in the LC site (12.8 %) though, ANOVA test did not showed a significant difference. On the other hand, the abundance of *Gammaproteobacteria* was slightly higher in the LC and HC sites (10.8 % and 10.4 %) than in the MC site (7.2 %) though, the difference was also not significant (Figure 4.7. A).

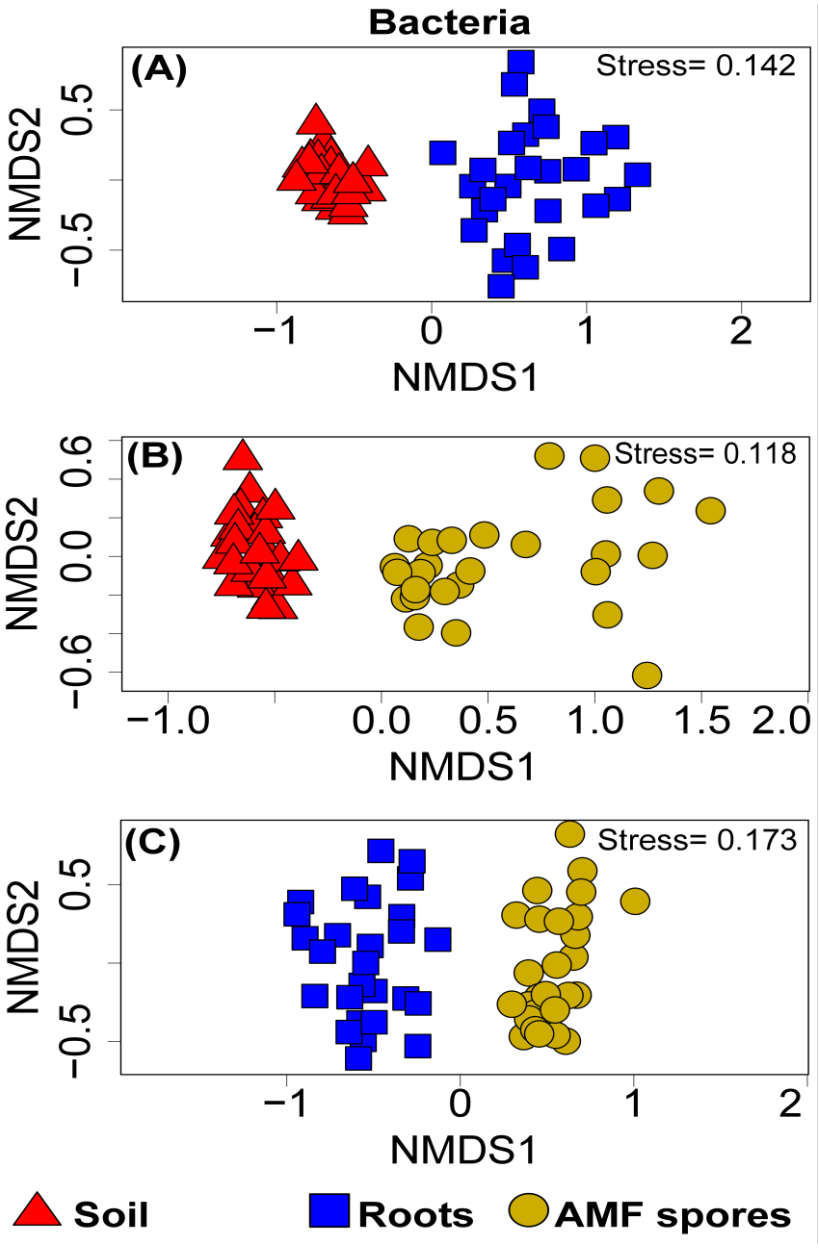


Figure 4.3. Non-metric multidimensional scaling (NMDS) showing the bacterial community compositions assignments across: (A) soil and roots (stress value = 0.14), (B) soil and AMF spores (stress value = 0.11) and (C) roots and AMF spores (stress value = 0.17).

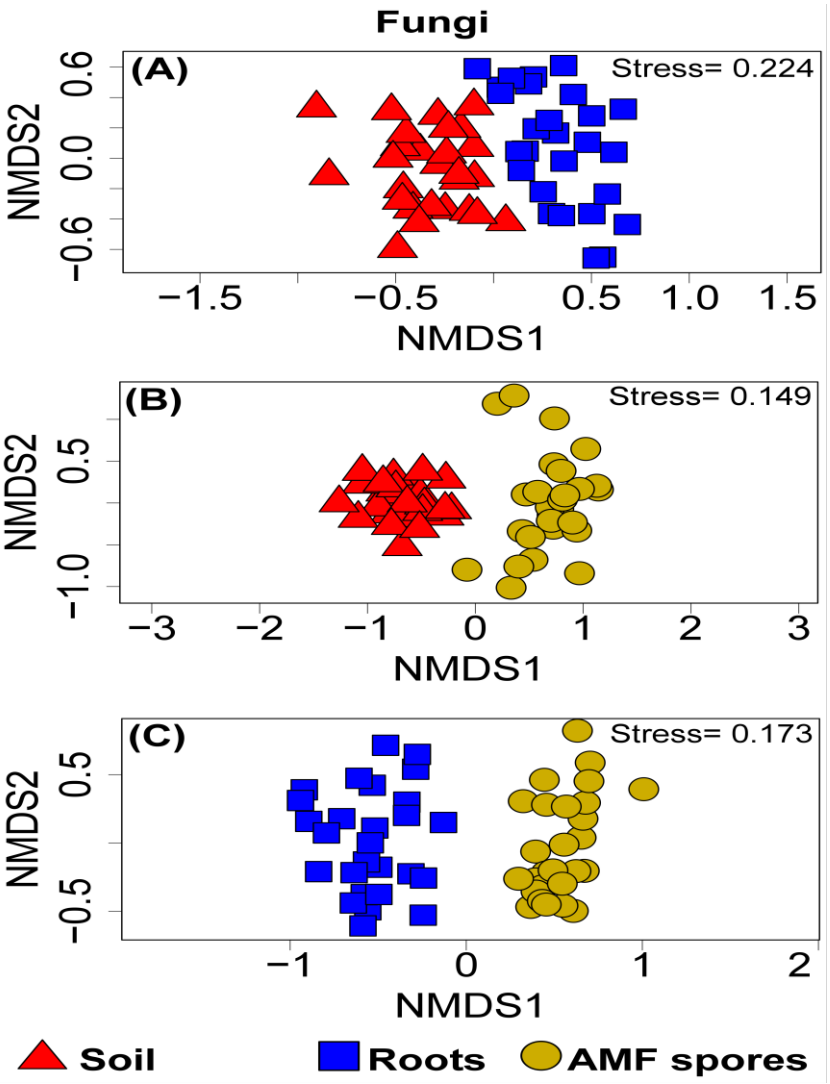


Figure 4.4. Non-metric multidimensional scaling (NMDS) showing the fungal community compositions assignments across: (A) soil and roots (stress value = 0.22), (B) soil and AMF spores (stress value = 0.14) and (C) roots and AMF spores (stress value = 0.17).

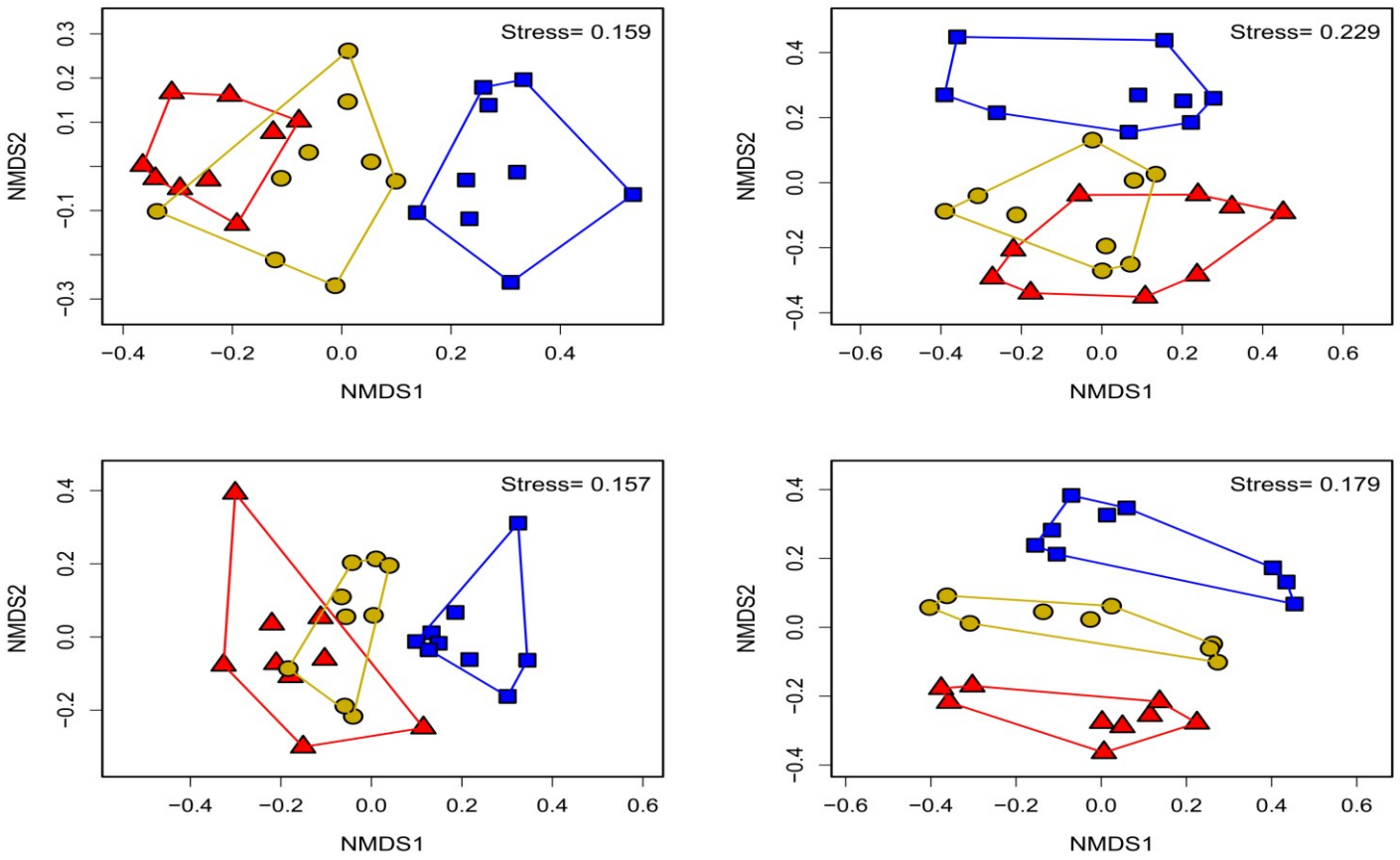


Figure 4.5. Non-metric multidimensional scaling (NMDS) showing the community composition assignments of: (A) soil bacteria (stress value = 0.15), (B) root bacteria (stress value = 0.22), (C) soil fungi (stress value = 0.15) and (D) root fungi across contamination concentrations (stress value = 0.17). PERMANOVA analysis showed significant effects of contamination levels on the community composition of soil bacteria, root bacteria, soil fungi and root fungi (n = 9, *P* = 0.010, 0.032, 0.003 and 0.006, respectively).

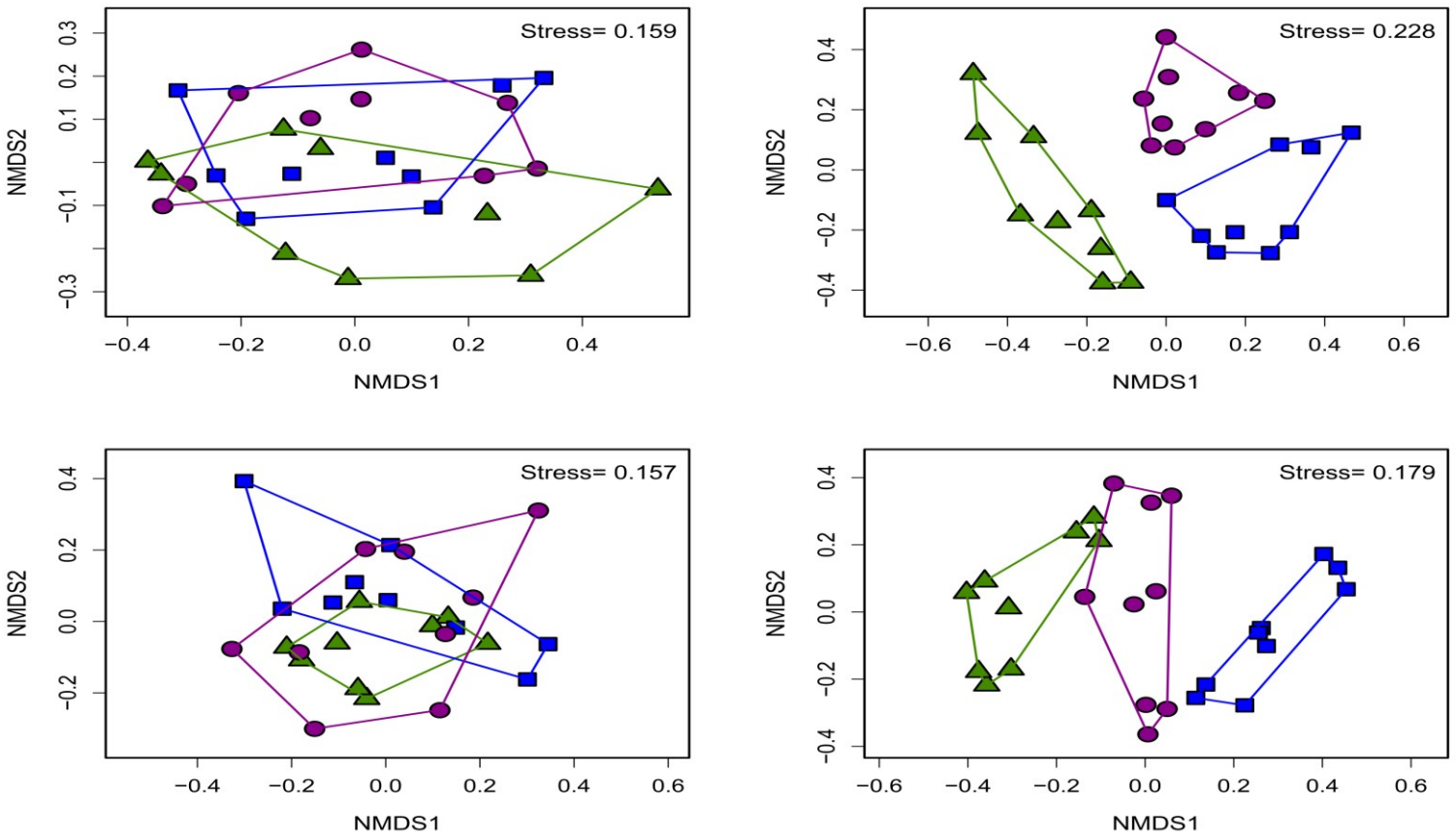


Figure 4.6. Non-metric multidimensional scaling (NMDS) showing the community compositions assignments of: (A) soil bacteria (stress value = 0.15), (B) root bacteria (stress value = 0.22), (C) soil fungi (stress value = 0.15) and (D) root fungi per plant species identity (stress value = 0.17). PERMANOVA analysis showed significant effects of plant species identity only on the community composition of root bacteria and root fungi (n = 9, *P* = 0.0001 and 0.0009, respectively).

Kruskal-Wallis tests performed on the root bacteria OTUs showed that the most abundant OTUs were affected by both contamination levels and plant species identity. Among the 30 most abundant OTUs, 10 were significantly affected by the contamination concentration, 13 by plant species and two by both contamination and plant species. Bacteria related to the classes *Alphaproteobacteria* (*Bradyrhizobium*, *Skermanella*, *Sphingobium*, *Hoeflea*, *Hyphomicrobium* and *Altererythrobacter*), *Actinobacteria* (*Streptomyces*, *Actinoplanes*, *Streptomyces* and *Lentzea*) *Betaproteobacteria* (*Ideonella*, *Duganella* and *Limnobacter*) and *Gammaproteobacteria* (*Rhizobacter*, *Steroidobacter* and *Pseudomonas*), had the most OTUs affected by the contamination levels or plant species (Table S4.2.).

When comparing the relative abundances of the root bacteria at class level across the contaminated sites, we observed that *Actinobacteria* were in higher proportions in the HC (20.4 %) and MC sites (22.5 %) than in the LC site (15.6 %) (ANOVA, $P = 0.041$). On the other hand, *Betaproteobacteria* was in higher abundance in the LC (13.9 %) and MC sites (14.1 %) than in the HC site (6.9 %) (ANOVA, $P = 0.003$).

When the relative abundances of root bacteria were compared across plant species, we observed that *Alphaproteobacteria* were clearly more abundant in *P. balsamifera* (38.3 %) and *L. europaeus* (33.8 %) than in *S. canadensis* (25.8 %) samples (ANOVA, $P < 0.001$). By contrast, the abundance of *Gammaproteobacteria* was higher in *S. canadensis* (24.2 %) than in *P. balsamifera* (10.1 %) and *L. europaeus* (8.4 %) samples (Kruskal-Wallis, $P = 0.004$). For *Betaproteobacteria*, its abundance was higher in *S. canadensis* and *L. europaeus* (12.6 % in both plant species roots), compared to *P. balsamifera* (9.6 %) but the *P* value was not significant (Figure 4.7. B).

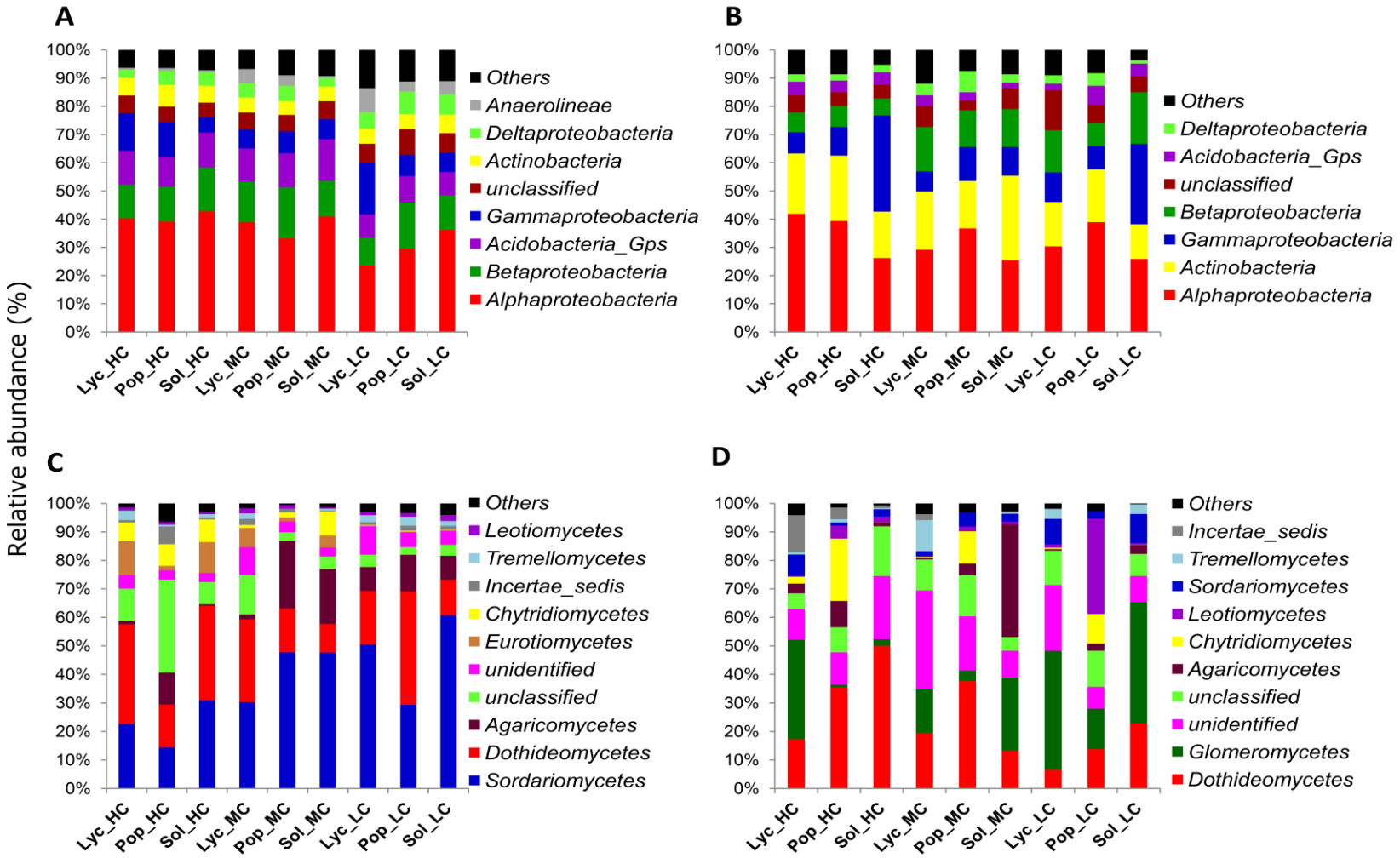


Figure 4.7. Relative abundances of major: (A) soil bacteria classes, (B) root bacteria classes, (C) soil fungi classes and (D) root fungi classes.

Sol_HC: *S. canadensis* in the highly contaminated site; Pop_HC: *P. balsamifera* in the highly contaminated site; Lyc_HC: *L. europaeus* in the highly contaminated site; Sol_MC: *S. canadensis* in the moderately contaminated site; Pop_MC: *P. balsamifera* in the moderately contaminated site; Lyc_MC: *L. europaeus* in the moderately contaminated site; Sol_LC: *S. canadensis* in the least contaminated site; Pop_LC: *P. balsamifera* in the least contaminated site; Lyc_LC: *L. europaeus* in the loweast contaminated site.

4.5.3. PHP and plant species identity effects on soil and root fungal diversity

ANOVA tests showed that there was no effect of contamination on the Shannon diversity indices of fungi in either soil ($P = 0.424$) or root ($P = 0.714$) datasets (Table 4.1 and Figure 4.2. A). However, there was a highly significant effect of plant species identity on the fungal diversity in roots ($P < 0.001$). Tukey's range test showed that the divergence in root fungal diversity has occurred in *L. europaeus*, which showed the highest diversity compared with *P. balsamifera* and *L. europaeus*. No effect of plant species identity on the fungal diversity was found in soil (Table 4.1 and Figure 4.2. B).

However, PERMANOVA analysis showed that the contamination had a significant effect on the community structure of fungi in both soil and root samples ($P = 0.003$ and $P = 0.006$, respectively), while plant species identity had a significant effect only in roots ($P = 0.003$) (Table 4.1). NMDS plots showed a clear separation between the community composition of the HC and LC sites, while the community of the MC site was intermediary in both soil and root datasets (Figure 4.5. C and D). NMDS plots across plant species showed differences in community structure only in roots, where a distinct grouping of the fungal communities was found between *L. europaeus* and *P. balsamifera*, with the community of *S. canadensis* being intermediate (Figure 4.6. C and D).

The Kruskal-Wallis tests confirmed that soil fungi were more affected by the contamination levels than by plants species identity, while a similar amount of root fungi were affected by either contamination and plants species (Table S4.2.). Among the 30 most abundant soil fungal OTUs, 13 were significantly affected by contamination, four by plants species, one by both, while the 12 remaining OTUs were not affected. The soil fungi OTU

which were significantly affected by contamination level or plants species identity belong mainly to the fungal classes *Sordariomycetes* (*Emericellopsis* sp. (OTU 5), *Lasiosphaeriaceae* (OTU 13) and *Fusarium* sp. (OTU 19)), *Agaricomycetes* (*Thelephoraceae* (OTU 13)), *Dothideomycetes* (*Leptosphaeria* sp. (OTU 11) and *Pycnidiophora* sp. (OTU 27)), *Eurotiomycetes* (*Penicillium* sp. (OTU 6)) and *Chytridiomycetes* (*Spizellomyces plurigibbosus* (OTU 8)). On the other hand, among the 30 most abundant root fungal OTUs, nine were affected by contamination, 10 by plants species, eight by both, while three OTUs were not affected. Most of the root fungi OTU which were significantly affected by contamination levels and/or plant species identity belong to the classes *Dothideomycetes* (*Leptosphaeria* sp. (OTU 1), *Pleosporales* sp. (OTU 15) and *Phoma herbarum* (OTU 7)), *Chytridiomycetes* (*Olpidium brassicae* (OTU 6) and *S. plurigibbosus* (OTU 2)), *Glomeromycetes* (*Claroideoglomus* (OTU 10) and *Entrophospora infrequens* (OTU 4)), *Leotiomycetes* (*Helotiales* (OTU 9)) and *Sordariomycetes* (*Fusarium sacchari* (OTU 22) and *Myrothecium* sp. (OTU 29)) and *Agaricomycetes* (Sebacinaceae (OTU 5)).

In soil, the fungal classes *Sordariomycetes* and *Agaricomycetes* were more abundant in the LC (46.9 % and 9.9 %, respectively) and MC sites (41.9 % and 14.9 %, respectively) than in the HC site (22.6 % and 4.3 %, respectively) (Kruskal-Wallis, $P = 0.02$ and 0.15). Contrarily, *Eurotiomycetes* and *Chytridiomycetes* were more abundant in the HC site (8.1 % and 7.3 %) than in MC (4.1 % and 3.7 %) and LC sites (0.1 % and 0.3 %) (Kruskal-Wallis, $P \leq 0.001$) (Figure 4.7. C and Table S4.2.). *Dothideomycetes* also showed a higher abundance in the HC (27.7 %) than in the MC and LC sites (18.1 % and 23.7 %) though, the Kruskal-Wallis rank test did not showed a significant difference. In roots, *Dothideomycetes* and *Chytridiomycetes* were more abundant in the HC site (34.3 % and 8.1 %) than in MC (23.4 %

and 3.9 %) and LC sites (14.4 % and 3.6 %) (Kruskal-Wallis, $P \leq 0.05$). Contrarily, *Glomeromycetes* and *Sordariomycetes* were more abundant in the LC site (32.8 % and 7.3 %) compared to MC (14.9 % and 3.2 %) and HC sites (12.7 % and 3.7 %) (Kruskal-Wallis, P = 0.02 and 0.04). When the abundances of root fungi were compared across plant species, we observed that the proportions of OTUs belonging to different fungal classes also varied between plant species identity. *Dothideomycetes* were more abundant in *P. balsamifera* (29 %) and *S. canadensis* (28.7 %) than in *L. europaeus* (14.4 %) (Kruskal-Wallis, P = 0.09), while *Glomeromycetes* and *Sordariomycetes* were more abundant in *L. europaeus* (30.7 % and 6.1 %, respectively) and *S. canadensis* (23.5 % and 5.2 %) than in *P. balsamifera* (6.3 % and 2.9 %, respectively) (Kruskal-Wallis, P = 0.004 and 0.14). *Agaricomycetes* were in higher proportions in *S. canadensis* (14.7 %) compared to *L. europaeus* (1.6 %) and *P. balsamifera* (5.3 %) (Kruskal-Wallis, P = 0.018). *Chytridiomycetes* and *Leotiomycetes* were more abundant in *P. balsamifera* (14.5 % and 13.2 %, respectively) than in *L. europaeus* (1 % and 0.6 %, respectively) and *S. canadensis* (0.04 % and 1.2 %, respectively) (Kruskal-Wallis, $P <$ 0.001) (Figure 4.7. D and Table S4.2.).

4.5.4. Soil and root microbial diversity versus AMF spore-associated microbial diversity

The comaparison of the soil and roots microbial communities with those identified in association to AMF in Iffis *et al.* (2016) study revealed that the community structure of AMF-associated microorganisms were significantly differed from the communities identified in the rhizosphere and endosphere of the same plants (PERMANOVA, $P < 0.001$). The NMDS plots showed a distinct grouping of soil and root microbial communities compared with AMF spore-associated microbial communities (Figure 4.3. C and D, and Figure 4.4. C and D). We

previously found that *Gammaproteobacteria* and *Betaproteobacteria* were the most dominant classes associated within AMF spores (their abundances were 49 % and 23 %, respectively), while the most dominant fungi belonged to the unclassified fungi (55 %), *Pezizomycetes* (13 %) and *Dothideomycetes* (13 %) (Iffis *et al.*, 2016). However, in the present study, *Alphaproteobacteria* was the most dominant bacterial class in both the soil and root datasets (36 % in soil and 33 % in roots), while *Sordariomycetes* was the most dominant fungal group in soil (37 %) and *Dothideomycetes* was the most dominant in roots (24 %). Here, *Gammaproteobacteria* represented only 9 % of OTUs in soil and 14 % in roots, while *Pezizomycetes* formed only 1.3 % of OTUs in roots and 0.8 % in soil (Figure S4.4.). Differences in community structures between the AMF spore-associated microbiomes and the rhizospheric and endophytic communities were also found at the genus level (Iffis *et al.*, 2016). Indeed, *Alphaproteobacteria* was represented mainly by the genera *Sphingomonas* in soils and *Bradyrhizobium* in roots, while this group was represented mainly by the genus *Caulobacter* in the AMF spore-associated microbiome (Iffis *et al.*, 2016). Similarly, *Betaproteobacteria* was represented mainly by the genera *Duganella* and *Janthinobacterium* in the spore microbiome, while they were formed mainly by unclassified *Betaproteobacteria* and *Caenimonas* in soil, and by *Duganella* in roots (Figure S4.2.). For fungi, *Septoria* was the most representative genus of *Dothideomycetes* associated with the AMF spores, while in soil and roots the *Dothideomycetes* was represented mainly by *Pleosporales* (Figure S4.3.).

4.6. Discussion

In rhizospheric soil, plant roots, bacteria and fungi form tripartite symbiosis or parasitic associations based on exchange of complex signaling dialogs and nutrient compounds by which each partner influences the other in order to avoid the different biotic and/or abiotic

stresses able to disrupt their life cycle. Therefore, the microbial community structures living in soil or in association with roots are intimately linked to the different exudates released in the rhizosphere (root and microbe exudates), soil composition and climatic conditions. In this study, we assessed the variation in bacterial and fungal diversity across PHP concentrations, plant species identity and biotopes (soil versus roots). Furthermore, since the soil and root samples used in this study are the same as those used previously in Iffis *et al.* (2016) study, the bacterial and fungal diversity found in the current study was compared to the AMF-associated microbial diversity found in the Iffis *et al.* (2016) study in order to verify the hypothesis that the microbial communities living in association with AMF spores are meticulously selected by the AMF exudates released in the mycosphere.

The comparison of microbial communities across PHP concentrations revealed that *Alphaproteobacteria* were favored in the high contaminated (HC) site, both in soil and in roots. *Actinobacteria* were also among the most dominant classes in the plant roots of the HC site. The high abundance of *Alphaproteobacteria* and *Actinobacteria* in the HC site may be related to their PHP tolerance and/or their ability to degrade PHP compounds. Several studies carried out on the microbial communities in PHP contaminated sites showed that *Alphaproteobacteria* and *Actinobacteria* were often found in higher abundances in the more highly organic contaminated soils (Greer *et al.*, 2010; Bell *et al.*, 2014; Yergeau *et al.*, 2014; Pagé *et al.*, 2015). Furthermore, it was reported that *Sphingomonas* (the most dominant *Alphaproteobacteria* in soil dataset) and *Bradyrhizobium* (the most dominant *Alphaproteobacteria* in root dataset), as well as *Streptomyces* (the most dominant *Actinobacteria* in soil and roots) are able to degrade a range of the recalcitrant polycyclic aromatic hydrocarbon (PAH) compounds, such as phenanthrene, pyrene and naphthalene

(Rentz *et al.*, 2008; Qu & Spain, 2011; Balachandran *et al.*, 2012; Bourguignon *et al.*, 2014). On the other hand, the presence of *Gammaproteobacteria* in similar abundances in the LC and HC sites, for both soil and root datasets, may be related to the large spectrum of activities of the species belonging to the *Gammaproteobacteria* class. For example, in PHP contaminated soils, *Pseudomonas* (the most dominant genus of *Gammaproteobacteria* in roots) was shown to degrade a range of PAH compounds such as phenanthrene, alkane and naphthalene (Ma *et al.*, 2006; Ní Chadhain & Zylstra, 2010; Sun *et al.*, 2014). In agricultural soils, *Pseudomonas* taxa are known as potential plant growth-promoting bacteria, where they are able to establish a symbiotic association with plant roots and play an important role in plant growth, nitrogen fixation and phosphate solubilization (Rodríguez & Fraga, 1999; Desnoues *et al.*, 2003; Sharma *et al.*, 2013). In the case of fungi, *Dothideomycetes* and *Chytridiomycetes* were the fungal classes which were found in higher abundance in the HC site. To our knowledge, there are no studies devoted to PHP tolerance or biodegradation abilities of *Chytridiomycetes.* Surprisingly, excepting the study of Iffis *et al.* (2016), none of the studies carried out nearby the basins of our study has found *Chytridiomycetes*, either in rhizospheric soils and sediments or in association with plant roots (Bell *et al.*, 2014; Stefani *et al.*, 2015; Bourdel *et al.*, 2016). However, all of these studies found *Dothideomycetes* among the most dominant fungal classes in the PHP contaminated sites. Furthermore, several studies reported that some species belonging to *Dothideomycetes* are able to tolerate or break down a range PHP compounds (Junghanns *et al.*, 2008; Ferrari *et al.*, 2011; Harms *et al.*, 2011; Stefani *et al.*, 2015). For example, *Alternaria* and *Cladosporium*, which were identified in both soil and root datasets, have been shown to degrade crude oil and a variety of its derivative products such as

phenanthrene, benzo[a]pyrene, fluoranthene, and anthracene (Giraud *et al.*, 2001; Potin *et al.*, 2004; Mohsenzadeh *et al.*, 2012; Ameen *et al.*, 2016).

Our resuts also showed that the OTU richness of bacteria and fungi were significantly decreased in root samples in comparison to the soil samples. Generally, the microbial diversity was shown to increase in rhizospheric soils compared to the different plant compartments (roots, stem or leaves) (Xu *et al.*, 2012; Turner *et al.*, 2013a; Edwards *et al.*, 2015). The increase of microbial richness in soils compared to roots may be related to the difference in environmental conditions and nutrient bioavailability in the two ecological niches (soil versus roots). Indeed, plant roots have a selective effect on both rhizospheric and endophytic microorganisms, however the selective effect is much higher in the endosphere (inside roots) owing to the complexity and specificity of plant-microbe interactions and plant immune system responses (Bais *et al.*, 2006; Hardoim *et al.*, 2008; Oldroyd, 2013). Generaly, before root colonization, plants and microorganisms engage in a complex chemical dialogue, and only the bacteria or fungi recognizing the signaling pathways are allowed to penetrate and colonize plant roots (Bertin *et al.*, 2003; Bais *et al.*, 2006; Oldroyd, 2013). Furthermore, once inside roots, the endophytes are subjected to stress caused by the new conditions and consequently, only the microbes able to adapt to the intraradical conditions can proliferate inside root compartments (Jones *et al.*, 2007; Parniske, 2008; Gaiero *et al.*, 2013; Brader *et al.*, 2014). For example, plant root infection by nitrogen fixing bacteria (eg. *Bradyrhizobium*) and arbuscular mycorrhizal fungi, which are found in higher proportion in root samples than in soil samples, is achieved through an exchange of complex chemical signaling between the plant roots and microbes. In the nitrogen-fixing bacterial symbiosis, plant roots release in the rhizosphere specific signaling compounds, composed mainly of flavonoids, which stimulate

nitrogen fixing bacteria to answer by producing a series of lipochitooligosaccharide compounds (nodulation factors), that are required to activate the rest of the symbiosis signaling pathway (Fisher & Long, 1992; Jones *et al.*, 2007). An analogue strategy to nitrogen-fixing bacteria infection was also described between plant roots and AMF. The signaling begins with root exudation of strigolactones in the rhizosphere. Perception of strigolactones by AMF stimulate the spores to answer by releasing other signaling compounds, so called "Myc factors", which trigger the symbiosis pathway (Parniske, 2008; Oldroyd, 2013). Contrary to the microorganisms living inside roots, in the rhizosphere the soil surrounding the roots is nutrient-rich. A large range of soil organic matter, as well as root exudates composed mainly of carbohydrates, amino acids and organic acids are present in the soil-root interface and stimulate the proliferation of the rhizosphere-living fungi and bacteria (Bertin *et al.*, 2003; Somers *et al.*, 2004; Bais *et al.*, 2006; Philippot *et al.*, 2013; Quiza *et al.*, 2015).

The comparison of microbial communities between soil ans roots showed that the proportions of OTUs belonging to some groups of fungi (eg: *Sordariomycetes* and *Glomeromycetes*) and bacteria (eg: *Gammaproteobacteria*, *Actinobacteria* and *Acidobacteria*) were different between rhizospheric soils and plant roots. In addition, at the genus rank, we found that the fungal and bacterial genera identified in rhizospheric soils were different from those identified in plant roots. Even if the root microbiome was considered as a community derived from the rhizospheric soil (Compant *et al.*, 2010; Gaiero *et al.*, 2013; Turner *et al.*, 2013a), several studies demonstrated that the microbial community composition (fungi and bacteria) in rhizospheric and bulk soils are different from those of plant roots (Smalla *et al.*, 2001; Xu *et al.*, 2012; Shakya *et al.*, 2013; Edwards *et al.*, 2015; Quiza *et al.*, 2015). Usually,

the same phyla or classes of microorganisms were found in soils and roots, but their abundances varied between the two biotopes. Moreover, differences in the taxonomic affiliations were often reported when the comparisons were carried out at genus or species rank (Gottel *et al.*, 2011; Turner *et al.*, 2013a; Turner *et al.*, 2013b; Edwards *et al.*, 2015). In our study, most of the bacteria and fungi identified in high proportions in roots were already known to be endophytic or obligatory biotrophic microorganisms, establishing a symbiotic, saprophytic or pathogenic associations with plants. For example, *Bradyrhizobium* and *Pseudomonas* (the most dominant *Alphaproteobacteria* and *Gammaproteobacteria* found in root dataset) are memebers of plant growth-promoting bacteria able to establish endosymbiotic associations with the roots of several plant species (Fisher & Long, 1992; Kneip *et al.*, 2007; Sharma *et al.*, 2013). Similarly, *Glomeromycetes* are known as obligate biotrophic fungi, which require a host plants for their growth and reproduction (Simon *et al.*, 1993; St-Arnaud & Vujanovic, 2007; Smith & Read, 2008).

On the other hand, the shifts in the community structures of AMF-associated bacteria and fungi across soils and roots observed in this study support the hypothesis that AMF select the microbial communities living in association with their spores and mycelia. As with plant roots, AMF may release carbon resources and other signaling molecules that render the surface of spores and mycelia favorable and then selective for the growth of specific microorganisms, as proposed previously (Roesti *et al.*, 2005; Bharadwaj *et al.*, 2011; Lecomte *et al.*, 2011; Agnolucci *et al.*, 2015; Iffis *et al.*, 2016). For example, Bharadwaj *et al.* (2011) conducted *in vitro* cultures of 10 AMF-associated bacteria isolates and they observed that the growth rate of the ten isolates were significantly increased by the addition to the culture medium a broth medium, in which AMF have been already cultured (Bharadwaj *et al.* (2011)

considered broth medium rich in AMF exudates). In another study, Roesti *et al.* (2005) performed four combinations of inoculations between two AMF species (*Glomus geosporum* and *Glomus constrictum*) and two host plant species (*Plantago lanceolata* and *Hieracium pilosella*), and they observed that the AMF spore identity affected the AMF-associated bacterial communities more than plant species identity did. However, little is known about AMF exudate compositions and their interactions with the soil microorganisms. Therefore, further investigations on this topic will be required to understand the different mechanisms by which AMF spores recruit their associated microorganisms.

4.7. Conclusion

The high throughput amplicon sequencing approach used in our study allowed us to identify variations in the bacterial and fungal communities in soils and roots across PH concentrations and plant species. Overall, we found that bacterial and fungal community structures associated to plant roots varied significantly across both PH concentrations and plant species identity. Whereas, bacterial and fungal communities found in soil were affected only by PH concentrations. Our results also showed that the OTU richness and community structures of bacteria and fungi were significantely differed between soil versus roots. Furthermore, comparisons between the AMF spore-associated microbiome described previously in Iffis *et al.* (2016) and the results of the present study showed that the microbial communities living in association with AMF spores differed from those found in the surrounding soil and roots.

4.8. Acknowledgements

This work was supported by the Genome Quebec and Genome Canada funded GenoRem Project which are gratefully acknowledged. We thank Karen Fisher-Favret for commenting on and editing the manuscript.

Chapitre 5 - Discussion générale et perspectives

5.1. Discussion

Dans les environnements contaminés, les plantes et les microorganismes du sol sont confrontés à des stress élevés et doivent développer des stratégies qui leur permettent de s'adapter à ces habitats hostiles. Plusieurs mécanismes ont été largement décrits dans la littérature, tels que la dégradation, la séquestration et la phyto-extraction des polluants qui sont présents dans la rhizosphère. De plus, des stratégies relatives aux modes de vie peuvent être mises en jeux par certains organismes en interagissant avec d'autres organismes pour établir des associations symbiotiques, saprophytiques ou parasitiques afin d'augmenter la tolérance aux conditions extrêmes causées par les différents polluants présents dans le sol. Les champignons mycorhiziens arbusculaires (CMA) font partie des microorganismes du sol qui jouent un rôle clé dans la phytoremédiation et ils sont capables d'interagir à la fois avec les racines de plantes et les microorganismes présents dans la rhizosphère.

Dans cette thèse de doctorat, j'ai mené une étude globale sur la biodiversité des bactéries et des champignons associés aux CMA dans des sites hautement contaminés par les hydrocarbures pétroliers (HP). Grâce aux techniques de clonage, de séquençage et de microscopie électronique à balayage, mes premières expériences ont abouti à la réalisation d'un portrait d'ensemble sur la biodiversité des bactéries associées aux propagules intra-racinaires des CMA extraites à partir des racines de *Solidago rugosa* qui poussent spontanément dans le site pollué de Varennes (chapitre 2). Mes résultats ont montré que malgré les concentrations en HP mesurées dans ce site ont atteint des niveaux extrêmes (voir Table S2.1), le niveau de diversité des CMA et des bactéries associées aux CMA a été élevé à

l'intérieur des racines de *S. rugosa*. Les observations microscopiques ont permis d'observer un taux de mycorhization des racines de *S. rugosa* pouvant atteindre 70 %, tandis que l'arbre phylogénétique obtenu à partir des séquences 18S de l'ARN ribosomique montre que les racines de *S. rugosa* sont colonisées par des espèces de CMA qui appartiennent au moins à cinq familles distinctes. Curieusement, parmi les séquences des clones 18S utilisés pour l'identification des CMA, six d'entre eux ont montré des similarités avec des espèces appartenant à la classe des Chytridiomycètes. Présentement, on connait peu de choses sur les relations entre les Chytridiomycètes et les CMA à l'exception de quelques études qui remontent aux années 1970 et 1980 et qui suggèrent que les Chytridiomycètes peuvent être soit des parasites, des saprotrophes des CMA ou même des vecteurs de bactéries (Ross & Ruttencutter, 1977; Schenck & Nicolson, 1977; Sylvia & Schenck, 1983; Tzean *et al.*, 1983; Paulitz & Menge, 1984). L'analyse des séquences 16S de l'ARN ribosomique a montré qu'il y a plusieurs espèces de bactéries qui sont associées aux propagules intra-racinaires des CMA. Le groupe des *mycorrhiza helper bacteria*, incluant des bactéries solubilisatrices de phosphate et fixatrices d'azote (e.g. les genres *Pseudomonas*, *Bosea*, *Brevundimonas*, *Bradyrhizobium*, et *Paenibacillus*), ainsi que des bactéries connues par leurs capacités à dégrader des hydrocarbures pétroliers (e.g. les genres *Pseudomonas*, *Sphingomonas*, *Massilia* et *Methylobacterium*) sont les taxons les plus dominants identifiés en association avec les propagules des CMA.

Les expériences du chapitre 3 m'ont permis d'investiguer en profondeur la biodiversité des CMA ainsi que celle des bactéries et champignons qui leurs sont associés dans les sites contaminés par les hydrocarbures pétroliers (HP). Pour ce faire, j'ai utilisé le séquençage à haut débit à partir des spores de CMA extraites de la rhizosphère de trois espèces de plantes

qui poussent spontanément dans trois bassins de décantation ayant des niveaux de contamination différents. Les résultats de cette expérience montrent que la biodiversité des microorganismes associés aux CMA est en réalité plus élevée que ce que les études précédentes avaient montré. Les espèces de champignons associées aux CMA identifiées dans ce chapitre appartiennent pratiquement à tous les embranchements des *Eumycota*, allant des *Chytridiomycota* jusqu'aux *Basidiomycota* avec une forte abondance des *Ascomycota*, tandis que la plupart des bactéries identifiées en association avec les spores des CMA sont des espèces telluriques appartenant dans la majorité des cas aux embranchements des *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, *Acidobacteria* et *Firmicutes*. La comparaison des communautés microbiennes associées aux CMA par rapport aux niveaux de contamination et aux affiliations taxonomiques des espèces de plantes montre que ces deux facteurs ont des effets significatifs sur la structure de ces communautés microbiennes. Il est important de noter que parmi les classes des microorganismes associés aux CMA identifiées avec des proportions élevées dans le bassin hautement contaminé, tels que *Actinobacteria* (bactéries), *Gammaproteobacteria* (bactéries) et *Dothideomycetes* (champignons), nombreuses contiennent des espèces connues pour leurs capacités à dégrader certains HP et en particulier les HAP. Les interactions entre les CMA et ces microorganismes dans des sites contaminés demeurent à ce jour méconnues, cependant certaines études ont démontré que la co-inoculation des CMA avec certaines espèces bactériennes peut augmenter significativement le rendement de la phytoremédiation (Alarcón *et al.*, 2008; Liu & Dalpé, 2009; Yu *et al.*, 2011; Dong *et al.*, 2014; Xun *et al.*, 2014). Les *Chytridiomycetes* font aussi partie des microorganismes associés aux CMA identifiés en proportions élevées dans le bassin hautement contaminé. Toutefois, comme discuté auparavant, on connait peu de choses sur les

interactions entre ces champignons et les CMA, et d'autant plus sur leurs rôles dans les processus de la phytoremédiation. Quant aux effets de l'affiliation taxonomique des plantes sur les structures des communautés microbiennes associées aux CMA, les exsudats racinaires jouent un rôle primordial dans la sélection des communautés microbiennes présentes dans les sols entourant les racines. Il a été démontré dans plusieurs études que les racines de plantes sécrètent des composés de signalisation, tels que les flavonoïdes, les strigolactones et d'autres molécules chimiques, qui sont capables d'établir des interactions très spécifiques avec certaines communautés microbiennes du sol (e.g. les interactions entre les légumineuses et les Rhizobiales) (Jones *et al.*, 2007; Turner *et al.*, 2013a; Edwards *et al.*, 2015; Rohrbacher & St-Arnaud, 2016). Grâce aux analyses de co-inertie, de partition de la variance et des corrélations de Spearman, j'ai observé aussi que certaines espèces de CMA peuvent avoir soit des corrélations positives ou négatives avec certains microorganismes associés à leurs spores. Cette observation m'a permis de supposer qu'en plus des effets des niveaux de contamination et de l'affiliation taxonomique des plantes, les CMA peuvent aussi jouer un rôle clé dans la sélection des communautés microbiennes associées à leurs spores et mycélia. Les mécanismes de recrutement des microorganismes par les CMA demeurent à ce jour méconnus, cependant des hypothèses suggèrent que les CMA sécrètent des exsudats mycéliens contenant des acides organiques, des sucres, des composés phénoliques et d'autres composés de signalisations destinés à stimuler et à mobiliser certains microorganismes d'intérêt présents dans le sol. (Roesti *et al.*, 2005; Bharadwaj *et al.*, 2011; Taktek *et al.*, 2015). En outre, étant donné que les CMA sont caractérisés par un polymorphisme génétique élevé entre les différentes espèces voire même entre les isolats de la même espèce (Hijri *et al.*, 1999; Redecker *et al.*, 1999; Hijri & Sanders, 2005; Croll *et al.*, 2008), il est fortement possible que la composition des exsudats

mycéliens varie d'une espèce de CMA à une autre, cependant cela reste à démontrer expérimentalement. De ce fait, j'ai formulé l'hypothèse que des espèces différentes de CMA peuvent avoir des affinités différentes vis-à-vis de la flore microbienne de la mycosphère.

Étant donné que les CMA sont des microorganismes telluriques, il est certain que les microbes recrutés par ceux-ci à la surface et à l'intérieur de leurs spores proviennent en grande majorité de la rhizosphère dans laquelle ils vivaient ou des racines des plantes hôtes qu'ils ont colonisées. Donc, il est pertinent de poser la question suivante: les communautés microbiennes associées aux CMA sont-elles différentes de celles de la rhizosphère et les racines de plantes?

Dans le chapitre 4 de mon projet de doctorat, j'ai tenté de répondre à cette question en étudiant les communautés microbiennes présentes dans les sols et les racines de plantes du même site qui a servi aux études précédentes (Varennes, QC), afin de les comparer avec celles en association avec les CMA. Pour ce faire, les échantillons des sols et des racines des plantes utilisées dans le chapitre 3 ont été récupérés et soumis au séquençage à haut débit pour étudier la biodiversité microbienne présente dans chacun des biotopes (les sols rhizosphériques et les racines). Mes résultats ont montrés malgré les CMA vivent à la fois dans les sols et à l'intérieur des racines de plantes, les communautés microbiennes présentent à la surface et à l'intérieur des spores des CMA sont différentes de celles identifiées dans les sols et racines. Même si j'ai observé pratiquement les mêmes classes de champignons et de bactéries dans les sols, les racines et les spores des CMA, leurs proportions étaient clairement différentes dans ces trois biotopes. De plus, quand la comparaison des communautés microbiennes des trois biotopes est réalisée au niveau des genres, j'ai observé que les affiliations taxonomiques des genres fongiques et bactériens observées en association avec les spores CMA sont dans la plus part des cas différentes de celles identifiées dans les racines et les sols. Cette différence entre

les communautés microbiennes associées aux CMA et celles associées aux sols et racines soutient l'hypothèse que les CMA sécrètent des exsudats mycéliens par lesquels ils sélectionnent la flore microbienne de leur mycosphère.

5.2. Les perspectives

Le présent projet de doctorat a permis une description détaillée des communautés microbiennes associées aux CMA présents dans les sites hautement contaminés par les HP. Les résultats obtenus montrent à quel point la biodiversité des microorganismes associés aux CMA est complexe dans le sol, même si celui-ci a atteint des niveaux extrêmes de pollution. Mon projet de doctorat n'a fait qu'effleurer la surface d'un monde aussi complexe que vaste qui peut concerner aussi bien les écosystèmes naturels que perturbés. La structure de ces communautés associées aux CMA est le résultat des échanges de complexes dialogues chimiques, dont on ne connait que peu de choses sur les molécules impliquées, les étapes des voies de signalisations ainsi que les facteurs qui peuvent influencer ces interactions. De ce fait, il serait donc nécessaire de se focaliser dans le futur sur l'étude de la structure des composés chimiques sécrétés par les CMA et les microorganismes qui sont leurs associés, notamment dans les milieux contaminés par les hydrocarbures pétroliers. Connaître la composition des exsudats mycéliens des CMA et savoir comment ceux-ci varient en fonction des plantes hôtes, des saisons, des nutriments (e.g. phosphore et azote) et des conditions physico-chimiques des sols, serait un atout considérable pour mieux comprendre les interactions entre les CMA et les microbes qui leurs sont associés. D'autre part, il est maintenant connu que les CMA jouent un rôle primordial dans les processus de phytoremédiation, cependant à ce jour on ne connait pas quels sont les mécanismes de décontamination que mettent en oeuvre ces champignons. Donc, il serait très pertinent de vérifier par des approches génomiques ou transcriptomiques la

présence des gènes codant pour des enzymes impliqués dans la dégradation, la chélation ou la translocation des HP chez les CMA. Pour finir, il y a des études plus récentes qui ont été réalisées *in vitro* et qui démontrent que le volume des réseaux mycéliens des champignons peut augmenter significativement la dissipation des hydrocarbures pétroliers dans les milieux de culture (Kohlmeier *et al.*, 2005; Banitz *et al.*, 2013; Schamfuß *et al.*, 2013). Ces études ont formulé l'hypothèse que les bactéries utilisent le réseau mycélien comme « autoroutes » (*fungal highways*) et se déplacent tout au long des filaments mycéliens, ayant ainsi une meilleure propagation et une meilleure exploration des interfaces de contact avec HP. Par la suite, ces bactéries se chargent de la dégradation des HP en composés secondaires assimilables et transportables par les hyphes des champignons « pipelines » (*fungal pipelines*). Dans le cas des processus de phytoremédiation par les CMA et bactéries, il me semble que ça serait très pertinent de vérifier en premier lieu cette hypothèse de *fungal highways* versus *fungal pipelines* pour comprendre les coopérations entre CMA et bactéries dans la dégradation des HP.

Bibliographie

Abdel Latef AAH, Chaoxing H. 2011. Effect of arbuscular mycorrhizal fungi on growth, mineral nutrition, antioxidant enzymes activity and fruit yield of tomato grown under salinity stress. *Scientia Horticulturae* **127**(3): 228-233.

Agnolucci M, Battini F, Cristani C, Giovannetti M. 2015. Diverse bacterial communities are recruited on spores of different arbuscular mycorrhizal fungal isolates. *Biology and Fertility of Soils* **51**(3): 379-389.

Alarcón A, Davies FT, Autenrieth RL, Zuberer DA. 2008. Arbuscular Mycorrhiza and Petroleum-Degrading Microorganisms Enhance Phytoremediation of Petroleum-Contaminated Soil. *International Journal of Phytoremediation* **10**(4): 251-263.

Ameen F, Moslem M, Hadi S, Al-Sabri AE. 2016. Biodegradation of diesel fuel hydrocarbons by mangrove fungi from Red Sea Coast of Saudi Arabia. *Saudi Journal of Biological Sciences* **23**(2): 211-218.

Angelard C, Colard A, Niculita-Hirzel H, Croll D, Sanders IR. 2010. Segregation in a mycorrhizal fungus alters rice growth and symbiosis-specific gene transcription. *Current Biology* **20**(13): 1216-1221.

Aranda E, Scervino JM, Godoy P, Reina R, Ocampo JA, Wittich R-M, García-Romera I. 2013. Role of arbuscular mycorrhizal fungus *Rhizophagus custos* in the dissipation of PAHs under root-organ culture conditions. *Environmental Pollution* **181**(0): 182-189.

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM. 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annual Review of Plant Biology* **57**(1): 233-266.

Balachandran C, Duraipandiyan V, Balakrishna K, Ignacimuthu S. 2012. Petroleum and polycyclic aromatic hydrocarbons (PAHs) degradation and naphthalene metabolism in *Streptomyces* sp. (ERI-CPDA-1) isolated from oil contaminated soil. *Bioresour Technol* **112**: 83-90.

Banitz T, Johst K, Wick LY, Schamfuß S, Harms H, Frank K. 2013. Highways versus pipelines: contributions of two fungal transport mechanisms to efficient bioremediation. *Environmental Microbiology Reports* **5**(2): 211-218.

Barea J-M, Pozo MJ, Azcón R, Azcón-Aguilar C. 2005. Microbial co-operation in the rhizosphere. *Journal of Experimental Botany* **56**(417): 1761-1778.

Battini F, Cristani C, Giovannetti M, Agnolucci M. 2016. Multifunctionality and diversity of culturable bacterial communities strictly associated with spores of the plant beneficial symbiont *Rhizophagus intraradices*. *Microbiological Research* **183**: 68-79.

Bedini S, Turrini A, Rigo C, Argese E, Giovannetti M. 2010. Molecular characterization and glomalin production of arbuscular mycorrhizal fungi colonizing a heavy metal polluted ash disposal island, downtown Venice. *Soil Biology and Biochemistry* **42**(5): 758-765.

Bell TH, Cloutier-Hurteau B, Al-Otaibi F, Turmel M-C, Yergeau E, Courchesne F, St-Arnaud M. 2015. Early rhizosphere microbiome composition is related to the growth and Zn uptake of willows introduced to a former landfill. *Environmental Microbiology* **17**(8): 3025-3038.

Bell TH, Hassan SE-D, Lauron-Moreau A, Al-Otaibi F, Hijri M, Yergeau E, St-Arnaud M. 2014. Linkage between bacterial and fungal rhizosphere communities in hydrocarbon-contaminated soils is related to plant phylogeny. *ISME J* **8**(2): 331-343.

Bertin C, Yang X, Weston L. 2003. The role of root exudates and allelochemicals in the rhizosphere. *Plant and Soil* **256**(1): 67-83.

Bharadwaj DP, Alström S, Lundquist P-O. 2011. Interactions among *Glomus irregulare*, arbuscular mycorrhizal spore-associated bacteria, and plant pathogens under in vitro conditions. *Mycorrhiza* **22**(6): 437-447.

Bianciotto V, Bandi C, Minerdi D, Sironi M, Tichy HV, Bonfante P. 1996. An obligately endosymbiotic mycorrhizal fungus itself harbors obligately intracellular bacteria. *Appl Environ Microbiol* **62**(8): 3005-3010.

Bianciotto V, Lumini E, Bonfante P, Vandamme P. 2003. *'Candidatus Glomeribacter gigasporarum'* gen. nov., sp. nov., an endosymbiont of arbuscular mycorrhizal fungi. *International Journal of Systematic and Evolutionary Microbiology* **53**(1): 121-124.

Bianciotto V, Lumini E, Lanfranco L, Minerdi D, Bonfante P, Perotto S. 2000. Detection and identification of bacterial endosymbionts in arbuscular mycorrhizal fungi belonging to the family *Gigasporaceae*. *Appl Environ Microbiol* **66**(10): 4503-4509.

Bieby Voijant T, Siti Rozaimah SA, Hassan B, Mushrifah I, Nurina A, Muhammad M. 2011. A review on heavy metals (As, Pb, and Hg) uptake by plants through phytoremediation. *International Journal of Chemical Engineering* **2011**.

Boer Wd, Folman LB, Summerbell RC, Boddy L. 2005. Living in a fungal world: impact of fungi on soil bacterial niche development. *FEMS Microbiology Reviews* **29**(4): 795-811.

Boffetta P, Jourenkova N, Gustavsson P. 1997. Cancer risk from occupational and environmental exposure to polycyclic aromatic hydrocarbons. *Cancer Causes & Control* **8**(3): 444-472.

Bonfante P, Anca IA. 2009. Plants, mycorrhizal fungi, and bacteria: a network of interactions. *Annu Rev Microbiol* **63**: 363-383.

Bourdel G, Roy-Bolduc A, St-Arnaud M, Hijri M. 2016. Concentration of petroleum-hydrocarbon contamination shapes fungal endophytic community structure in plant roots. *Frontiers in Microbiology* **7**.

Bourguignon N, Isaac P, Alvarez H, Amoroso MJ, Ferrero MA. 2014. Enhanced polyaromatic hydrocarbon degradation by adapted cultures of actinomycete strains. *Journal of Basic Microbiology* **54**(12): 1288-1294.

Bozzola JJ, Russell LD. 1992. *Electron microscopy : principles and techniques for biologists*. Boston Jones and Bartlett Publishers.

Brader G, Compant S, Mitter B, Trognitz F, Sessitsch A. 2014. Metabolic potential of endophytic bacteria. *Current Opinion in Biotechnology* **27**: 30-37.

Cabral L, Soares CRFS, Giachini AJ, Siqueira JO. 2015. Arbuscular mycorrhizal fungi in phytoremediation of contaminated areas by trace elements: mechanisms and major benefits of their applications. *World Journal of Microbiology and Biotechnology* **31**(11): 1655-1664.

Chao A, Jost L. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* **93**(12): 2533-2547.

Chen B, Kan H. 2008. Air pollution and population health: a global challenge. *Environmental Health and Preventive Medicine* **13**(2): 94-101.

Cho EJ, Lee DJ, Wee CD, Kim HL, Cheong YH, Cho JS, Sohn BK. 2009. Effects of AMF inoculation on growth of *Panax ginseng* C.A. Meyer seedlings and on soil structures in mycorrhizosphere. *Scientia Horticulturae* **122**(4): 633-637.

Compant S, Clément C, Sessitsch A. 2010. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. *Soil Biology and Biochemistry* **42**(5): 669-678.

Croll D, Wille L, Gamper HA, Mathimaran N, Lammers PJ, Corradi N, Sanders IR. 2008. Genetic diversity and host plant preferences revealed by simple sequence repeat and mitochondrial markers in a population of the arbuscular mycorrhizal fungus Glomus intraradices. *New Phytologist* **178**(3): 672-687.

Cruz AF, Ishii T. 2012. Arbuscular mycorrhizal fungal spores host bacteria that affect nutrient biodynamics and biocontrol of soil-borne plant pathogens. *Biology Open* **1**(1): 52-57.

Dennis JJ, Zylstra GJ. 2004. Complete sequence and genetic organization of pDTG1, the 83 kilobase naphthalene degradation plasmid from *Pseudomonas putida* strain NCIB 9816-4. *Journal of Molecular Biology* **341**(3): 753-768.

Desjardins D, Nissim WG, Pitre FE, Naud A, Labrecque M. 2014. Distribution patterns of spontaneous vegetation and pollution at a former decantation basin in southern Québec, Canada. *Ecological Engineering* **64**(0): 385-390.

Desnoues N, Lin M, Guo X, Ma L, Carreño-Lopez R, Elmerich C. 2003. Nitrogen fixation genetics and regulation in a Pseudomonas stutzeri strain associated with rice. *Microbiology* **149**(8): 2251-2262.

Dolédec S, Chessel D. 1994. Co-inertia analysis: an alternative method for studying species–environment relationships. *Freshwater Biology* **31**(3): 277-294.

Dong R, Gu L, Guo C, Xun F, Liu J. 2014. Effect of PGPR *Serratia marcescens* BC-3 and AMF *Glomus intraradices* on phytoremediation of petroleum contaminated soil. *Ecotoxicology* **23**(4): 674-680.

Dray S, xe, phane, Chessel D, Thioulouse J. 2003. Co-inertia analysis and the linking of ecological data tables. *Ecology* **84**(11): 3078-3089.

Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, Eisen JA, Sundaresan V. 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences* **112**(8): E911-E920.

Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H, Hugenholtz P. 2010. Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J* **4**(5): 642-647.

Ferrari BC, Zhang C, van Dorst J. 2011. Recovering Greater Fungal Diversity from Pristine and Diesel Fuel Contaminated Sub-Antarctic Soil Through Cultivation Using Both a High and a Low Nutrient Media Approach. *Frontiers in Microbiology* **2**: 217.

Fisher RF, Long SR. 1992. Rhizobium-plant signal exchange. *Nature* **357**(6380): 655-660.

Fortin JA, Plenchette C, Piché Y 2008. Les mycorhizes : La nouvelle révolution verte: Editions Multimondes.

Frey-Klett P, Garbaye J, Tarkka M. 2007. The mycorrhiza helper bacteria revisited. *New Phytologist* **176**(1): 22-36.

Gaiero JR, McCall CA, Thompson KA, Day NJ, Best AS, Dunfield KE. 2013. Inside the root microbiome: Bacterial root endophytes and plant growth promotion. *American Journal of Botany* **100**(9): 1738-1750.

Gao Y, Cao X, Kang F, Cheng Z. 2011a. PAHs pass through the cell wall and partition into organelles of arbuscular mycorrhizal roots of ryegrass All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. *J. Environ. Qual.* **40**(2): 653-656.

Gao Y, Cheng Z, Ling W, Huang J. 2010. Arbuscular mycorrhizal fungal hyphae contribute to the uptake of polycyclic aromatic hydrocarbons by plant roots. *Bioresour Technol* **101**(18): 6895-6901.

Gao Y, Li Q, Ling W, Zhu X. 2011b. Arbuscular mycorrhizal phytoremediation of soils contaminated with phenanthrene and pyrene. *Journal of Hazardous Materials* **185**(2–3): 703-709.

Garbaye J. 1994. Helper bacteria - a new dimension to the mycorrhizal symbiosis. *New Phytologist* **128**(2): 197-210.

Ghignone S, Salvioli A, Anca I, Lumini E, Ortu G, Petiti L, Cruveiller S, Bianciotto V, Piffanelli P, Lanfranco L, Bonfante P. 2012. The genome of the obligate endobacterium of an AM fungus reveals an interphylum network of nutritional interactions. *ISME J* **6**(1): 136-145.

Gihring TM, Green SJ, Schadt CW. 2012. Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. *Environmental Microbiology* **14**(2): 285-290.

Giovannetti M, Mosse B. 1980. An evaluation of techniques for measuring vesicular arbuscular mycorrhizal infection in roots. *New Phytologist* **84**(3): 489-500.

Giraud F, Guiraud P, Kadri M, Blake G, Steiman R. 2001. Biodegradation of anthracene and fluoranthene by fungi isolated from an experimental constructed wetland for wastewater treatment. *Water Research* **35**(17): 4126-4136.

Gottel NR, Castro HF, Kerley M, Yang Z, Pelletier DA, Podar M, Karpinets T, Uberbacher E, Tuskan GA, Vilgalys R, Doktycz MJ, Schadt CW. 2011. Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types. *Appl Environ Microbiol* **77**(17): 5934-5944.

Greer CW, Whyte LG, Niederberger TD 2010. Microbial communities in hydrocarbon-contaminated temperate, tropical, alpine, and polar soils. In: Timmis KN ed. *Handbook of Hydrocarbon and Lipid Microbiology*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2313-2328.

Gremion F, Chatzinotas A, Kaufmann K, von Sigler W, Harms H. 2004. Impacts of heavy metal contamination and phytoremediation on a microbial community during a twelve-month microcosm experiment. *FEMS Microbiology Ecology* **48**(2): 273-283.

Hardoim PR, van Overbeek LS, Elsas JDv. 2008. Properties of bacterial endophytes and their proposed role in plant growth. *Trends in Microbiology* **16**(10): 463-471.

Harms H, Schlosser D, Wick LY. 2011. Untapped potential: exploiting fungi in bioremediation of hazardous chemicals. *Nat Rev Micro* **9**(3): 177-192.

Hassan SE-D, Bell TH, Stefani FOP, Denis D, Hijri M, St-Arnaud M. 2014. Contrasting the community structure of arbuscular mycorrhizal fungi from hydrocarbon-

contaminated and uncontaminated soils following willow (*Salix* spp. L.) planting. *PLoS ONE* **9**(7): e102838.

Hassan SE-D, Boon E, St-Arnaud M, Hijri M. 2011. Molecular biodiversity of arbuscular mycorrhizal fungi in trace metal-polluted soils. *Molecular Ecology* **20**(16): 3469-3483.

Hassan SE-D, Hijri M, St-Arnaud M. 2013. Effect of arbuscular mycorrhizal fungi on trace metal uptake by sunflower plants grown on cadmium contaminated soil. *New Biotechnology* **30**(6): 780-787.

Henner P, Schiavon M, Morel J-L, Lichtfouse E. 1997. Polycyclic aromatic hydrocarbon (PAH) occurrence and remediation methods. *Analusis* **25**(9-10): M56-M59.

Hernández-Ortega HA, Alarcón A, Ferrera-Cerrato R, Zavaleta-Mancera HA, López-Delgado HA, Mendoza-López MR. 2012. Arbuscular mycorrhizal fungi on growth, nutrient status, and total antioxidant activity of *Melilotus albus* during phytoremediation of a diesel-contaminated substrate. *Journal of Environmental Management* **95, Supplement**: S319-S324.

Hijri M, Hosny M, van Tuinen D, Dulieu H. 1999. Intraspecific ITS polymorphism in *Scutellospora castanea*(*Glomales*, *Zygomycota*) is structured within multinucleate spores. *Fungal Genetics and Biology* **26**(2): 141-151.

Hijri M, Redecker D, Petetot JAM-C, Voigt K, Wöstemeyer J, Sanders IR. 2002. Identification and isolation of two *Ascomycete* fungi from spores of the arbuscular mycorrhizal fungus *Scutellospora castanea*. *Appl Environ Microbiol* **68**(9): 4567-4573.

Hijri M, Sanders IR. 2005. Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei. *Nature* **433**(7022): 160-163.

Hildebrandt U, Ouziad F, Marner F-J, Bothe H. 2006. The bacterium *Paenibacillus validus* stimulates growth of the arbuscular mycorrhizal fungus Glomus intraradices up to the formation of fertile spores. *FEMS Microbiol Lett* **254**(2): 258-267.

Hu J, Lin X, Wang J, Cui X, Dai J, Chu H, Zhang J. 2010. Arbuscular mycorrhizal fungus enhances P acquisition of wheat (*Triticum aestivum* L.) in a sandy loam soil with long-term inorganic fertilization regime. *Applied Microbiology and Biotechnology* **88**(3): 781-787.

Iffis B, St-Arnaud M, Hijri M. 2014. Bacteria associated with arbuscular mycorrhizal fungi within roots of plants growing in a soil highly contaminated with aliphatic and aromatic petroleum hydrocarbons. *FEMS Microbiol Lett* **358**: 44–54.

Iffis B, St-Arnaud M, Hijri M. 2016. Petroleum hydrocarbon contamination, plant identity and arbuscular mycorrhizal fungal community determine assemblages of the AMF spore-associated microbes. *Environmental Microbiology*: n/a-n/a.

Ismail Y, McCormick S, Hijri M. 2011. A fungal symbiont of plant-roots modulates mycotoxin gene expression in the pathogen *Fusarium sambucinum*. *PLoS ONE* **6**(3): e17990.

Ismail Y, McCormick S, Hijri M. 2013. The arbuscular mycorrhizal fungus, *Glomus irregulare*, controls the mycotoxin production of *Fusarium sambucinum* in the pathogenesis of potato. *FEMS Microbiol Lett* **348**(1): 46-51.

Jargeat P, Cosseau C, Ola'h B, Jauneau A, Bonfante P, Batut J, Bécard G. 2004. Isolation, Free-Living Capacities, and Genome Structure of "*Candidatus Glomeribacter gigasporarum*" the Endocellular Bacterium of the Mycorrhizal Fungus *Gigaspora margarita*. *Journal of Bacteriology* **186**(20): 6876-6884.

Jing YD, He ZL, Yang XE. 2007. Role of soil rhizobacteria in phytoremediation of heavy metal contaminated soils. *J Zhejiang Univ Sci B* **8**(3): 192-207.
Jobard M, Rasconi S, Sime-Ngando T. 2010. Diversity and functions of microscopic fungi: a missing component in pelagic food webs. *Aquatic Sciences* **72**(3): 255-268.
Joner EJ, Johansen A, Loibner AP, Dela Cruz MA, Szolar OHJ, Portal JM, Leyval C. 2001. Rhizosphere effects on microbial community structure and dissipation and toxicity of polycyclic aromatic hydrocarbons (PAHs) in spiked soil. *Environ Sci Technol* **35**(13): 2773-2777.
Joner EJ, Leyval C. 2003. Rhizosphere gradients of polycyclic aromatic hydrocarbon (PAH) dissipation in two industrial soils and the impact of arbuscular mycorrhiza. *Environ Sci Technol* **37**(11): 2371-2375.
Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC. 2007. How rhizobial symbionts invade plants: the *Sinorhizobium*-Medicago model. *Nat Rev Micro* **5**(8): 619-633.
Junghanns C, Krauss G, Schlosser D. 2008. Potential of aquatic fungi derived from diverse freshwater environments to decolourise synthetic azo and anthraquinone dyes. *Bioresour Technol* **99**(5): 1225-1235.
Karagiannidis N, Thomidis T, Lazari D, Panou-Filotheou E, Karagiannidou C. 2011. Effect of three Greek arbuscular mycorrhizal fungi in improving the growth, nutrient concentration, and production of essential oils of oregano and mint plants. *Scientia Horticulturae* **129**(2): 329-334.
Kivlin SN, Hawkes CV, Treseder KK. 2011. Global diversity and distribution of arbuscular mycorrhizal fungi. *Soil Biology and Biochemistry* **43**(11): 2294-2303.
Klabi R, Bell TH, Hamel C, Iwaasa A, Schellenberg M, Raies A, St-Arnaud M. 2014. Plant assemblage composition and soil P concentration differentially affect communities of AM and total fungi in a semi-arid grassland. *FEMS Microbiology Ecology* **91**: 1-13.
Kneip C, Lockhart P, VoSZ C, Maier U-G. 2007. Nitrogen fixation in eukaryotes - new models for symbiosis. *BMC Evolutionary Biology* **7**(1): 55.
Kohlmeier S, Smits THM, Ford RM, Keel C, Harms H, Wick LY. 2005. Taking the fungal highway: mobilization of pollutant-degrading bacteria by fungi. *Environ Sci Technol* **39**(12): 4640-4646.
Kok M, Oldenhuis R, van der Linden MP, Raatjes P, Kingma J, van Lelyveld PH, Witholt B. 1989. The *Pseudomonas oleovoran*s alkane hydroxylase gene. Sequence and expression. *J Biol Chem* **264**(10): 5435-5441.
Lace B, Genre A, Woo S, Faccio A, Lorito M, Bonfante P. 2015. Gate crashing arbuscular mycorrhizas: in vivo imaging shows the extensive colonization of both symbionts by *Trichoderma atroviride*. *Environmental Microbiology Reports* **7**(1): 64-77.
Lecomte J, St-Arnaud M, Hijri M. 2011. Isolation and identification of soil bacteria growing at the expense of arbuscular mycorrhizal fungi. *FEMS Microbiology Letters* **317**(1): 43-51.
Lee J, Lee S, Young JPW. 2008. Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiology Ecology* **65**(2): 339-349.
Legendre P, Legendre L. 2012. *Numerical Ecology*: Elsevier Science BV, Amsterdam.
Leyval C, Joner EJ, del Val C, Haselwandter K. 2002. Potential of arbuscular mycorrhizal fungi for bioremediation. *Mycorrhizal Technology in Agriculture*: 175-186.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**(13): 1658-1659.
Liu A, Dalpé Y. 2009. Reduction in soil polycyclic aromatic hydrocarbons by arbuscular mycorrhizal leek plants. *International Journal of Phytoremediation* **11**(1): 39-52.
Liu SL, Luo YM, Cao ZH, Wu LH, Ding KQ, Christie P. 2004. Degradation of benzo[a]pyrene in soil with arbuscular mycorrhizal alfalfa. *Environmental Geochemistry and Health* **26**(2-3): 285-293.
Lloyd-Macgilp SA, Chambers SM, Dodd JC, Fitter AH, Walker C, Young JPW. 1996. Diversity of the ribosomal internal transcribed spacers within and among isolates of *Glomus mosseae* and related mycorrhizal fungi. *New Phytologist* **133**(1): 103-111.
Long L, Zhu H, Yao Q, Ai Y. 2008. Analysis of bacterial communities associated with spores of Gigaspora margarita and Gigaspora rosea. *Plant and Soil* **310**(1): 1-9.
Long LK, Yao Q, Guo J, Yang RH, Huang YH, Zhu HH. 2010. Molecular community analysis of arbuscular mycorrhizal fungi associated with five selected plant species from heavy metal polluted soils. *European Journal of Soil Biology* **46**(5): 288-294.
Lumini E, Orgiazzi A, Borriello R, Bonfante P, Bianciotto V. 2010. Disclosing arbuscular mycorrhizal fungal biodiversity in soil through a land-use gradient using a pyrosequencing approach. *Environmental Microbiology* **12**(8): 2165-2179.
Ma Y, Rajkumar M, Zhang C, Freitas H. 2016. Beneficial role of bacterial endophytes in heavy metal phytoremediation. *Journal of Environmental Management* **174**: 14-25.
Ma Y, Wang L, Shao Z. 2006. *Pseudomonas*, the dominant polycyclic aromatic hydrocarbon-degrading bacteria isolated from antarctic soils and the role of large plasmids in horizontal gene transfer. *Environmental Microbiology* **8**(3): 455-465.
MacDonald RM, Chandler M, Mosse B. 1982. The occurrence of bacterium-like organelles in vesicular-arbuscular mycorrhizal fungi. *New Phytol* **90**: 659-663.
Mahanty B, Pakshirajan K, Dasu VV. 2011. Understanding the complexity and strategic evolution in PAH remediation research. *Critical Reviews in Environmental Science and Technology* **41**(19): 1697-1746.
Marleau J, Dalpe Y, St-Arnaud M, Hijri M. 2011. Spore development and nuclear inheritance in arbuscular mycorrhizal fungi. *BMC Evolutionary Biology* **11**.
Marschner P, Timonen S 2006. Bacterial Community Composition and Activity in Rhizosphere of Roots Colonized by Arbuscular Mycorrhizal Fungi. In: Mukerji KG, Manoharachary C, Singh J eds. *Microbial Activity in the Rhizoshere*: Springer Berlin Heidelberg, 139-154.
McAloose D, Newton AL. 2009. Wildlife cancer: a conservation perspective. *Nat Rev Cancer* **9**(7): 517-526.
McGrath SP, Zhao FJ, Lombi E. 2001. Plant and rhizosphere processes involved in phytoremediation of metal-contaminated soils. *Plant and Soil* **232**(1): 207-214.
Mirabal Alonso L, Kleiner D, Ortega E. 2008. Spores of the mycorrhizal fungus *Glomus mosseae* host yeasts that solubilize phosphate and accumulate polyphosphates. *Mycorrhiza* **18**(4): 197-204.
Miransari M. 2011. Interactions between arbuscular mycorrhizal fungi and soil bacteria. *Applied Microbiology and Biotechnology* **89**(4): 917-930.
Mishra V, Gupta A, Kaur P, Singh S, Singh N, Gehlot P, Singh J. 2016. Synergistic effects of arbuscular mycorrhizal fungi and plant growth promoting rhizobacteria in

bioremediation of iron contaminated soils. *International Journal of Phytoremediation* **18**(7): 697-703.

Mohamed EF, Awad G, Andriantsiferana C, El-Diwany AI. 2015. Biofiltration technology for the removal of toluene from polluted air using Streptomyces griseus. *Environmental Technology*: 1-11.

Mohsenzadeh F, Chehregani Rad A, Akbari M. 2012. Evaluation of oil removal efficiency and enzymatic activity in some fungal strains for bioremediation of petroleum-polluted soils. *Iranian Journal of Environmental Health Science & Engineering* **9**(1): 26-26.

Mosse B. 1970. Honey-coloured, sessile Endogone spores: II. Changes in fine structure during spore development. *Archiv für Mikrobiologie* **74**(2): 129-145.

Naumann M, Schusler A, Bonfante P. 2010. The obligate endobacteria of arbuscular mycorrhizal fungi are ancient heritable components related to the *Mollicutes*. *ISME J* **4**(7): 862-871.

Ní Chadhain SM, Zylstra GJ 2010. Functional gene diversity, biogeography, dynamics. In: Timmis K ed. *Handbook of Hydrocarbon and Lipid Microbiology*: Springer, Berlin, Heidelberg, 2413-2422.

Oehl F, Sieverding E, Palenzuela J, Ineichen K, Alves da Silva G. 2011. Advances in Glomeromycota taxonomy and classification. *IMA Fungus* **2**(2): 191-199.

Oldroyd GED. 2013. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat Rev Micro* **11**(4): 252-263.

Ondov BD, Bergman NH, Phillippy AM. 2011. Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* **12**: 385-385.

Öpik M, Vanatoa A, Vanatoa E, Moora M, Davison J, Kalwij JM, Reier Ü, Zobel M. 2010. The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (*Glomeromycota*). *New Phytologist* **188**(1): 223-241.

Pagé AP, Yergeau É, Greer CW. 2015. *Salix purpurea* stimulates the expression of specific bacterial xenobiotic degradation genes in a soil contaminated with hydrocarbons. *PLoS ONE* **10**(7): e0132062.

Parniske M. 2008. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nat Rev Micro* **6**(10): 763-775.

Paulitz TC, Menge JA. 1984. Is *Spizellomyces punctatum* a parasite or saprophyte of vesicular-arbuscular mycorrhizal fungi?. *Mycologia* **76**(1): 99-107.

Peuke AD, Rennenberg H. 2005. Phytoremediation. *Molecular biology, requirements for application, environmental protection, public attention and feasibility* **6**(6): 497-501.

Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH. 2013. Going back to the roots: the microbial ecology of the rhizosphere. *Nat Rev Micro* **11**(11): 789-799.

Pilon-Smits E. 2005. Phytoremediation. *Annual Review of Plant Biology* **56**(1): 15-39.

Poirier MC. 2004. Chemical-induced DNA damage and human cancer risk. *Nat Rev Cancer* **4**(8): 630-637.

Posada RH, Franco LA, Ramos C, Plazas LS, Suárez JC, Álvarez F. 2008. Effect of physical, chemical and environmental characteristics on arbuscular mycorrhizal fungi in Brachiaria decumbens (Stapf) pastures. *Journal of Applied Microbiology* **104**(1): 132-140.

Potin O, Veignie E, Rafin C. 2004. Biodegradation of polycyclic aromatic hydrocarbons (PAHs) by *Cladosporium sphaerospermum* isolated from an aged PAH contaminated soil. *FEMS Microbiology Ecology* **51**(1): 71-78.

Qu Y, Spain JC. 2011. Molecular and biochemical characterization of the 5-nitroanthranilic acid degradation pathway in *Bradyrhizobium* sp. strain JS329. *Journal of Bacteriology* **193**(12): 3057-3063.

Quiza L, St-Arnaud M, Yergeau E. 2015. Harnessing phytomicrobiome signalling for rhizosphere microbiome engineering. *Frontiers in Plant Science* **6**.

Redecker D, Hijri M, Dulieu H, Sanders IR. 1999. Phylogenetic analysis of a dataset of fungal 5.8S rDNA sequences shows that highly divergent copies of internal transcribed spacers reported from *Scutellospora castanea* are of ascomycete origin. *Fungal Genetics and Biology* **28**(3): 238-244.

Redecker D, Kodner R, Graham LE. 2000. Glomalean fungi from the Ordovician. *Science* **289**(5486): 1920-1921.

Redecker D, Schüßler A, Stockinger H, Stürmer S, Morton J, Walker C. 2013. An evidence-based consensus for the classification of arbuscular mycorrhizal fungi (Glomeromycota). *Mycorrhiza* **23**(7): 515-531.

Rentz JA, Alvarez PJJ, Schnoor JL. 2008. Benzo[a]pyrene degradation by *Sphingomonas yanoikuyae* JAR02. *Environmental Pollution* **151**(3): 669-677.

Roche 2009. Using multiplex identifier (MID) adaptors for the GS FLX titanium chemistry–extended MID set. *Technical Bulletin: Genome Sequencer FLX System TCB no 005-2009 Roche*. Branchburg,NJ.

Rodríguez H, Fraga R. 1999. Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnology Advances* **17**(4–5): 319-339.

Roesti D, Ineichen K, Braissant O, Redecker D, Wiemken A, Aragno M. 2005. Bacteria associated with spores of the arbuscular mycorrhizal fungi *Glomus geosporum* and *Glomus constrictum*. *Applied and Environmental Microbiology* **71**(11): 6673-6679.

Rohrbacher F, St-Arnaud M. 2016. Root exudation: the ecological driver of hydrocarbon rhizoremediation. *Agronomy* **6**(1): 19.

Ross J, Ruttencutter R. 1977. Population dynamics of two vesicular arbuscular endomycorrhizal fungi and the role of hyperparasitic fungi. *Phytopathology* **67**: 490–496.

Salt DE, Blaylock M, Kumar NPBA, Dushenkov V, Ensley BD, Chet I, Raskin I. 1995. Phytoremediation: A novel strategy for the removal of toxic metals from the environment using plants. *Nat Biotech* **13**(5): 468-474.

Salt DE, Smith RD, Raskin I. 1998. Phytoremediation. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**(1): 643-668.

Salvioli A, Ghignone S, Novero M, Navazio L, Venice F, Bagnaresi P, Bonfante P. 2016. Symbiosis with an endobacterium increases the fitness of a mycorrhizal fungus, raising its bioenergetic potential. *ISME J* **10**(1): 130-144.

Samanta SK, Singh OV, Jain RK. 2002. Polycyclic aromatic hydrocarbons: environmental pollution and bioremediation. *Trends in Biotechnology* **20**(6): 243-248.

Sanders IR, Alt M, Groppe K, Boller T, Wiemken A. 1995. Identification of ribosomal DNA polymorphisms among and within spores of the *Glomales*: application to studies on the genetic diversity of arbuscular mycorrhizal fungal communities. *New Phytologist* **130**(3): 419-427.

Santos-González JC, Finlay RD, Tehler A. 2007. Seasonal Dynamics of Arbuscular Mycorrhizal Fungal Communities in Roots in a Seminatural Grassland. *Applied and Environmental Microbiology* **73**(17): 5613-5623.

Schamfuß S, Neu TR, van der Meer JR, Tecon R, Harms H, Wick LY. 2013. Impact of mycelia on the accessibility of fluorene to PAH-degrading bacteria. *Environ Sci Technol* **47**(13): 6908-6915.

Schenck NC, Nicolson TH. 1977. A zoosporic fungus occurring on species of *Gigaspora margarita* and other vesicular-arbuscular mycorrhizal fungi. *Mycologia* **69**(5): 1049-1053.

Scheublin TR, Sanders IR, Keel C, van der Meer JR. 2010. Characterisation of microbial communities colonising the hyphal surfaces of arbuscular mycorrhizal fungi. *ISME J* **4**(6): 752-763.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**(23): 7537-7541.

Schouteden N, De Waele D, Panis B, Vos CM. 2015. Arbuscular mycorrhizal fungi for the biocontrol of plant-parasitic nematodes: a review of the mechanisms involved. *Frontiers in Microbiology* **6**.

Schüβler A, Schwarzott D, Walker C. 2001. A new fungal phylum, the Glomeromycota: phylogeny and evolution. *Mycological Research* **105**(12): 1413-1421.

Shakya M, Gottel N, Castro H, Yang ZK, Gunter L, Labbé J, Muchero W, Bonito G, Vilgalys R, Tuskan G, Podar M, Schadt CW. 2013. A multifactor analysis of fungal and bacterial community structure in the root microbiome of mature *Populus deltoides* trees. *PLoS ONE* **8**(10): e76382.

Sharma S, Sayyed R, Trivedi M, Gobi T. 2013. Phosphate solubilizing microbes: sustainable approach for managing phosphorus deficiency in agricultural soils. *SpringerPlus* **2**(1): 587.

Simon L, Bousquet J, Levesque RC, Lalonde M. 1993. Origin and diversification of endomycorrhizal fungi and coincidence with vascular land plants. *Nature* **363**(6424): 67-69.

Simon L, Lalonde M, Bruns TD. 1992. Specific amplification of 18S fungal ribosomal genes from vesicular-arbuscular endomycorrhizal fungi colonizing roots. *Appl Environ Microbiol* **58**(1): 291-295.

Smalla K, Wieland G, Buchner A, Zock A, Parzy J, Kaiser S, Roskot N, Heuer H, Berg G. 2001. Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: Plant-dependent enrichment and seasonal shifts revealed. *Appl Environ Microbiol* **67**(10): 4742-4751.

Smith S, Read D. 2008. *Mycorrhizal Symbiosis*: Academic Press,London.

Somers E, Vanderleyden J, Srinivasan M. 2004. Rhizosphere Bacterial Signalling: A Love Parade Beneath Our Feet. *Critical Reviews in Microbiology* **30**(4): 205-240.

Song Y, Chen D, Lu K, Sun Z, Zeng R. 2015. Enhanced tomato disease resistance primed by arbuscular mycorrhizal fungus. *Frontiers in Plant Science* **6**.

St-Arnaud M, Elsen A 2005. Interaction of arbuscular mycorrhizal fungi with soil-borne pathogens and non-pathogenic rhizosphere micro-organisms. In: Declerck S, Fortin JA,

Strullu D-G eds. *In Vitro Culture of Mycorrhizas*: Springer, Berlin, Heidelberg, 217-231.

St-Arnaud M, Vujanovic V 2007. Effect of the arbuscular mycorrhizal symbiosis on plant diseases and pests. In: Chantal H, Christian P eds. *Mycorrhizae in Crop Production*. Binghampton, New York, USA: Haworth Food & Agricultural Products Press, 67-122.

Stefani FOP, Bell TH, Marchand C, de la Providencia IE, El Yassimi A, St-Arnaud M, Hijri M. 2015. Culture-dependent and -independent methods capture different microbial community fractions in hydrocarbon-contaminated soils. *PLoS ONE* **10**(6): e0128272.

Strobel G, Daisy B. 2003. Bioprospecting for microbial endophytes and their natural products. *Microbiology and Molecular Biology Reviews* **67**(4): 491-502.

Sun K, Liu J, Gao Y, Jin L, Gu Y, Wang W. 2014. Isolation, plant colonization potential, and phenanthrene degradation performance of the endophytic bacterium *Pseudomonas* sp. Ph6-gfp. *scientific Reports* **4**: 5462.

Sundram S, Meon S, Seman IA, Othman R. 2011. Symbiotic interaction of endophytic bacteria with arbuscular mycorrhizal fungi and its antagonistic effect on *Ganoderma boninense*. *The Journal of Microbiology* **49**(4): 551-557.

Sylvia DM, Schenck NC. 1983. Soil fungicides for controlling Chytridiaceous mycoparasites of *Gigaspora margarita* and *Glomus fasciculatum*. *Appl Environ Microbiol* **45**(4): 1306-1309.

Taktek S, Trépanier M, Servin PM, St-Arnaud M, Piché Y, Fortin JA, Antoun H. 2015. Trapping of phosphate solubilizing bacteria on hyphae of the arbuscular mycorrhizal fungus *Rhizophagus irregularis* DAOM 197198. *Soil Biology and Biochemistry* **90**: 1-9.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28**(10): 2731-2739.

Tarkka M, Frey-Klett P 2008. Mycorrhiza helper bacteria. In: Varma A ed. *Mycorrhiza*. Berlin, Heidelberg: Springer Berlin Heidelberg, 113-132.

Taylor TN, Remy W, Hass H, Kerp H. 1995. Fossil arbuscular mycorrhizae from the early Devonian. *Mycologia* **87**: 560-573.

Teng Y, Luo Y, Sun X, Tu C, Xu L, Liu W, Li Z, Christie P. 2010. Influence of arbuscular mycorrhiza and *Rhizobium* on phytoremediation by alfalfa of an agricultural soil contaminated with weathered pcbs: A field study. *International Journal of Phytoremediation* **12**(5): 516-533.

Thijs S, Sillen W, Rineau F, Weyens N, Vangronsveld J. 2016. Towards an enhanced understanding of plant-microbiome interactions to improve phytoremediation: Engineering the metaorganism. *Frontiers in Microbiology* **7**.

Toju H, Sato H, Tanabe AS. 2014. Diversity and spatial structure of belowground plant–fungal symbiosis in a mixed subtropical forest of ectomycorrhizal and arbuscular mycorrhizal plants. *PLoS ONE* **9**(1): e86566.

Turner TR, James EK, Poole PS. 2013a. The plant microbiome. *Genome Biology* **14**(6): 209-209.

Turner TR, Ramakrishnan K, Walshaw J, Heavens D, Alston M, Swarbreck D, Osbourn A, Grant A, Poole PS. 2013b. Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *ISME J* **7**(12): 2248-2258.

Tyagi M, da Fonseca MMR, de Carvalho CCCR. 2011. Bioaugmentation and biostimulation strategies to improve the effectiveness of bioremediation processes. *Biodegradation* **22**(2): 231-241.

Tzean S, Chu C, Su H. 1983. Spiroplasmalike organisms in a vesicular-arbuscular mycorrhizal fungus and its mycoparasite. *Phytopathology* **73**: 989–991.

Vallino M, Massa N, Lumini E, Bianciotto V, Berta G, Bonfante P. 2006. Assessment of arbuscular mycorrhizal fungal diversity in roots of *Solidago gigantea* growing in a polluted soil in Northern Italy. *Environmental Microbiology* **8**(6): 971-983.

Van Aken B, Yoon JM, Schnoor JL. 2004. Biodegradation of nitro-substituted explosives 2,4,6-trinitrotoluene, hexahydro-1,3,5-trinitro-1,3,5-triazine, and octahydro-1,3,5,7-tetranitro-1,3,5-tetrazocine by a phytosymbiotic *Methylobacterium* sp. associated with *Poplar* tissues (*Populus deltoides* × nigra DN34). *Applied and Environmental Microbiology* **70**(1): 508-517.

Van Geel M, Busschaert P, Honnay O, Lievens B. 2014. Evaluation of six primer pairs targeting the nuclear rRNA operon for characterization of arbuscular mycorrhizal fungal (AMF) communities using 454 pyrosequencing. *Journal of Microbiological Methods* **106**: 93-100.

Vannini C, Carpentieri A, Salvioli A, Novero M, Marsoni M, Testa L, de Pinto MC, Amoresano A, Ortolani F, Bracale M, Bonfante P. 2016. An interdomain network: the endobacterium of a mycorrhizal fungus promotes antioxidative responses in both fungal and plant hosts. *New Phytologist*: n/a-n/a.

Volante A, Lingua G, Cesaro P, Cresta A, Puppo M, Ariati L, Berta G. 2005. Influence of three species of arbuscular mycorrhizal fungi on the persistence of aromatic hydrocarbons in contaminated substrates. *Mycorrhiza* **16**(1): 43-50.

Walker C, Mize CW, McNabb Jr. HS. 1982. Populations of endogonaceous fungi at two locations in central Iowa. *Canadian Journal of Botany* **60**(12): 2518-2529.

Wu N, Huang H, Zhang S, Zhu Y-G, Christie P, Zhang Y. 2009. Phenanthrene uptake by *Medicago sativa* L. under the influence of an arbuscular mycorrhizal fungus. *Environmental Pollution* **157**(5): 1613-1618.

Xu L, Ravnskov S, Larsen J, Nicolaisen M. 2012. Linking fungal communities in roots, rhizosphere, and soil to the health status of Pisum sativum. *FEMS Microbiology Ecology* **82**(3): 736-745.

Xun F, Xie B, Liu S, Guo C. 2014. Effect of plant growth-promoting bacteria (PGPR) and arbuscular mycorrhizal fungi (AMF) inoculation on oats in saline-alkali soil contaminated by petroleum to enhance phytoremediation. *Environmental Science and Pollution Research* **22**(1): 598-608.

Xun F, Xie B, Liu S, Guo C. 2015. Effect of plant growth-promoting bacteria (PGPR) and arbuscular mycorrhizal fungi (AMF) inoculation on oats in saline-alkali soil contaminated by petroleum to enhance phytoremediation. *Environmental Science and Pollution Research* **22**(1): 598-608.

Yasmeen T, Hameed S, Tariq M, Ali S. 2012. Significance of arbuscular mycorrhizal and bacterial symbionts in a tripartite association with Vigna radiata. *Acta Physiologiae Plantarum* **34**(4): 1519-1528.

Yergeau E, Sanschagrin S, Maynard C, St-Arnaud M, Greer CW. 2014. Microbial expression profiles in the rhizosphere of willows depend on soil contamination. *ISME J* **8**(2): 344-358.

Yu XZ, Wu SC, Wu FY, Wong MH. 2011. Enhanced dissipation of PAHs from soil using mycorrhizal ryegrass and PAH-degrading bacteria. *Journal of Hazardous Materials* **186**(2–3): 1206-1217.

Zhang W, Wang H, Zhang R, Yu X-Z, Qian P-Y, Wong MH. 2010. Bacterial communities in PAH contaminated soils at an electronic-waste processing center in China. *Ecotoxicology* **19**(1): 96-104.

Zhou HW, Guo CL, Wong YS, Tam NF. 2006. Genetic diversity of dioxygenase genes in polycyclic aromatic hydrocarbon-degrading bacteria isolated from mangrove sediments. *FEMS Microbiol Lett* **262**(2): 148-157.

Annexes

1- Supporting information (Chapter 2)

Bacteria associated with arbuscular mycorrhizal fungi within roots of plants growing in a soil highly contaminated with aliphatic and aromatic petroleum hydrocarbons

Table S2.1. Concentrations of polycyclic aromatic hydrocarbons (PAH) and alkanes (C10–C50) in the sediments where *Solidago rugosa* plants were collected.

PAHs	Accepted limit values[1]	Concentrations (mg/kg)
Acenaphthene	100	**760**
Anthracene	100	**340**
Benzo anthracene	10	**21**
Chrysene	10	**23**
Fluorene	100	**710**
Phenanthrene	50	**2700**
Pyrene	100	**150**
2-Methylnaphthalene	10	**160**
1-Methylnaphthalene	10	**320**
1,3-Dimethylnaphthalene	10	**390**
2,3,5-Trimethylnaphthalene	10	**150**
Alkanes (C10-C50)	3500	**41000**

[1]Values corresponding to the maximum limit accepted for contamination of industrial areas by the Sustainable Development, Environment, Wildlife and Parks Department of the Province of Québec, Canada (http://www.mddefp.gouv.qc.ca/sol/terrains/politique/annexe_2_tableau_1.htm).

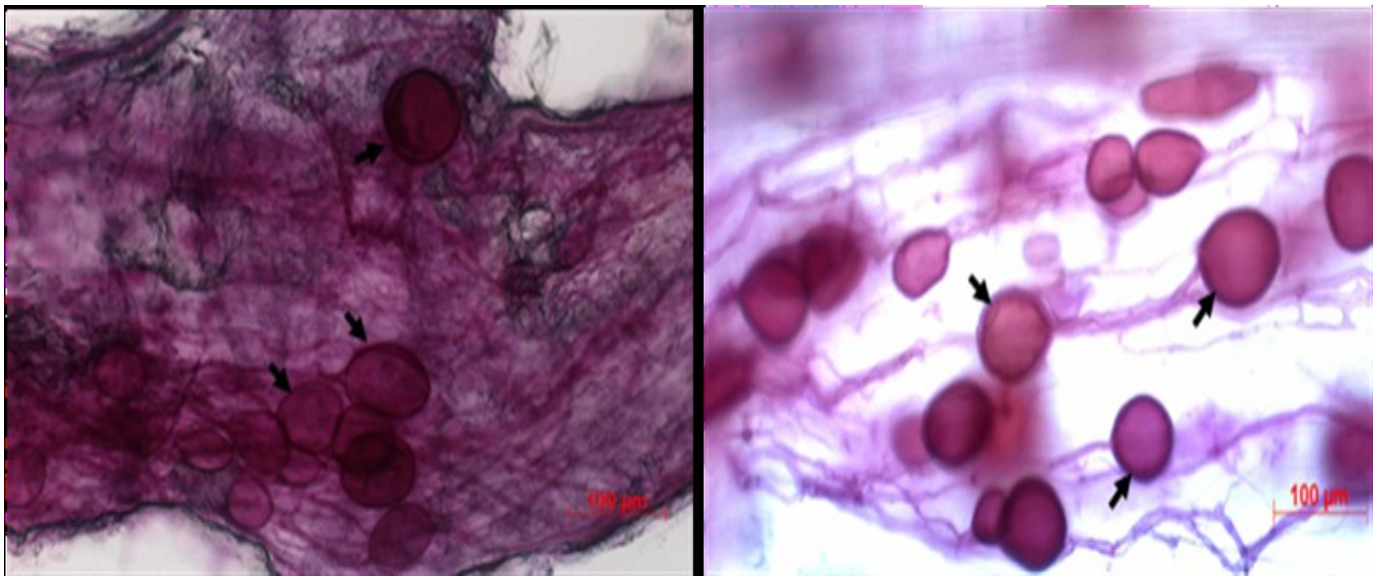


Figure S2.1. Root of Solidago rugosa stained using fuchsine, showing a high intensity of mycorrhizal colonization.

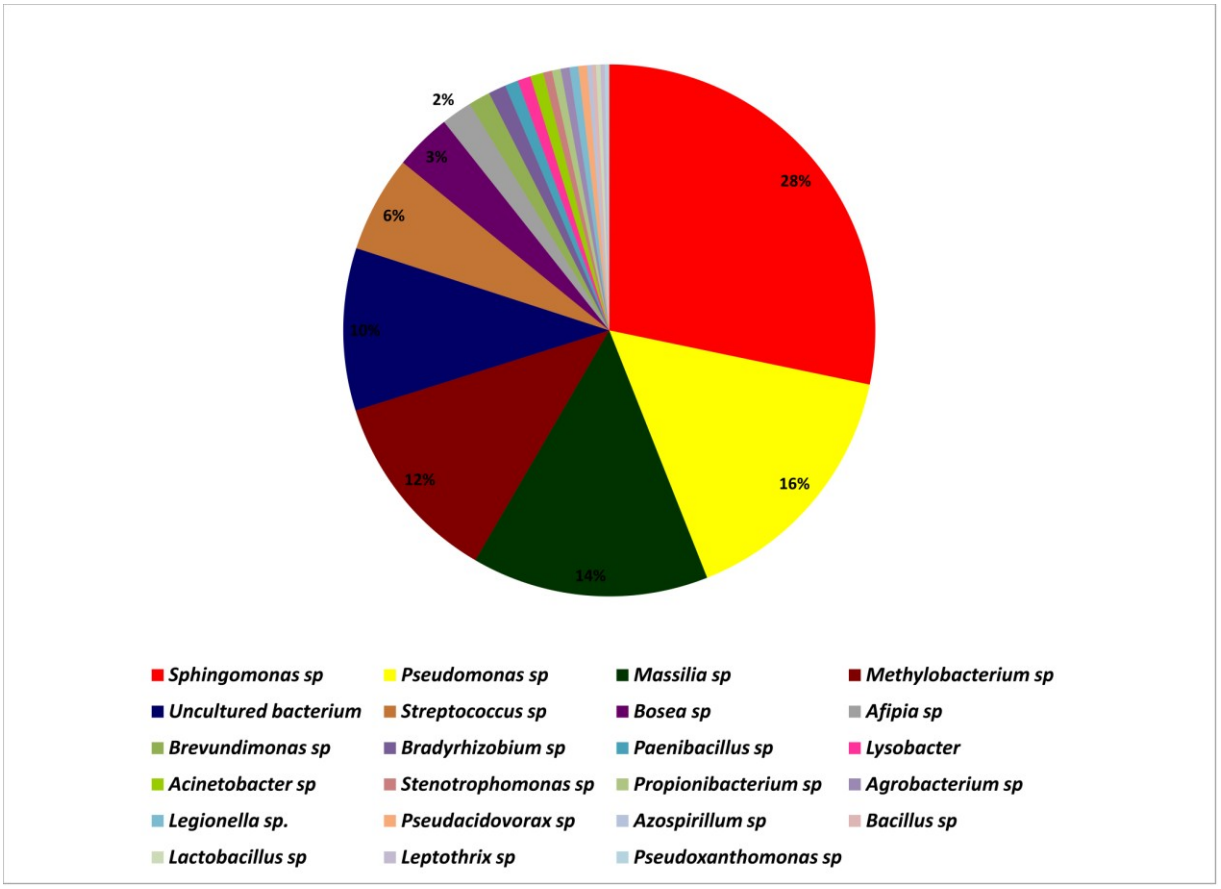


Figure S2.2. Proportion of the different bacterial genera identified in AMF propagules and Solidago rugosa roots. Only proportions over 2 % are showed.

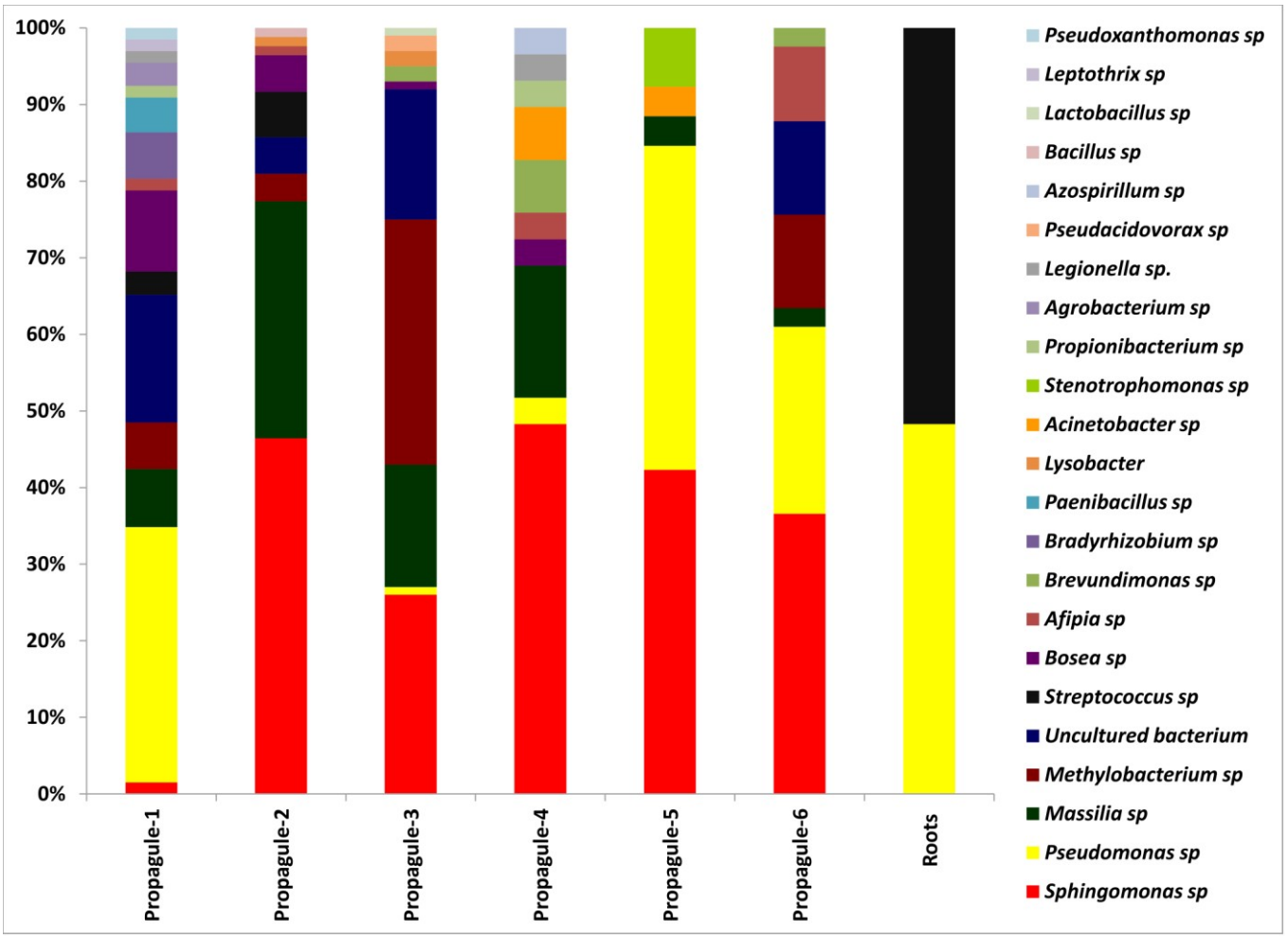


Figure S2.3. Stacked histogram showing the percentage of the different bacterial genera in each of the six AMF propagules and roots.

2- Supporting information (Chapter 3)

Petroleum hydrocarbon contamination, plant identity and arbuscular mycorrhizal fungal community determine assemblages of the AMF spore-associated microbes

Table S3.1. AMF virtual taxa of the 18S rRNA gene dataset after sub-sampling. OTUs sequences were compared with MaarjAM database.
This supporting table is available as Table S1 in the website of Environmental Microbiology Journal (http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13438/full).

Table S3.2. OTUs of AMF ITS dataset after sub-sampling.
This supporting table is available as Table S2 in the website of Environmental Microbiology Journal (http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13438/full).

Table S3.3. OTUs of non AMF fungi dataset after sub-sampling.
This supporting table is available as Table S3 in the website of Environmental Microbiology Journal (http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13438/full).

Table S3.4. OTUs of bacteria 16S rRNA genes dataset after sub-sampling.
This supporting table is available as Table S4 in the website of Environmental Microbiology Journal (http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13438/full).

Table S3.5. Observed richness, Chao1 estimator values and Good's coverage values for each individual sample across the different datasets. LY: *Lycopus europaeus*, PO: *Populus balsamifera*, SO: *Solidago canadensis*, HC: high contamination, MC: moderate contamination, LC: low contamination.

	AMF (18S dataset)			AMF (ITS dataset)			Fungi (without AMF)			Bacteria		
Samples	**sobs**	**chao**	**coverage**	**sobs**	**chao**	**coverage**	**sobs**	**chao**	**coverage**	**sobs**	**chao**	**coverage**
LY1-HC	7.00	7.50	1.00	5.00	5.00	1.00	4.00	5.00	0.96	138.00	246.97	0.92
LY1-LC	7.00	7.00	1.00	7.00	10.00	0.99	10.00	10.33	0.96	206.00	330.90	0.89
LY1-MC	2.00	2.00	1.00	6.00	7.00	0.99	6.00	7.00	0.96	217.00	304.00	0.91
LY2-HC	9.00	9.33	1.00	12.00	12.00	1.00	3.00	3.00	0.98	117.00	255.32	0.93
LY2-LC	5.00	5.00	1.00	8.00	9.00	0.99	16.00	20.20	0.85	140.00	248.37	0.92
LY2-MC	8.00	11.00	1.00	9.00	15.00	0.98	10.00	17.50	0.87	298.00	458.16	0.85
LY3-HC	11.00	11.33	1.00	8.00	11.00	0.99	5.00	8.00	0.94	177.00	319.50	0.91
LY3-LC	7.00	8.00	1.00	9.00	9.75	0.99	7.00	13.00	0.91	198.00	343.41	0.89
LY3-MC	11.00	11.75	1.00	11.00	14.00	0.98	14.00	16.50	0.89	169.00	319.19	0.90
PO1-HC	7.00	8.00	1.00	11.00	14.33	0.98	7.00	17.00	0.89	49.00	124.43	0.97
PO1-LC	9.00	10.50	1.00	8.00	11.00	0.99	6.00	7.00	0.96	97.00	155.33	0.95
PO1-MC	7.00	7.00	1.00	10.00	16.00	0.98	9.00	9.33	0.96	194.00	338.76	0.90
PO2-HC	5.00	5.00	1.00	6.00	7.50	0.99	5.00	5.50	0.96	54.00	95.33	0.97
PO2-LC	15.00	18.00	0.99	10.00	10.00	1.00	6.00	12.00	0.91	108.00	242.00	0.93
PO2-MC	9.00	12.00	1.00	9.00	15.00	0.98	4.00	4.00	0.98	212.00	352.14	0.89
PO3-HC	4.00	4.00	1.00	6.00	6.50	0.99	8.00	9.50	0.94	259.00	409.96	0.87
PO3-LC	4.00	4.00	1.00	4.00	4.00	1.00	5.00	5.00	0.98	174.00	268.22	0.92
PO3-MC	10.00	12.00	0.99	15.00	22.50	0.98	7.00	9.00	0.91	174.00	257.13	0.91
SO1-HC	8.00	14.00	0.99	8.00	8.00	1.00	9.00	12.00	0.94	74.00	113.67	0.97
SO1-LC	7.00	7.00	1.00	12.00	13.50	0.98	4.00	4.00	0.98	30.00	60.60	0.98
SO1-MC	4.00	4.00	1.00	7.00	8.00	0.99	8.00	14.00	0.91	155.00	238.42	0.92
SO2-HC	9.00	15.00	0.99	7.00	7.00	1.00	6.00	6.50	0.96	65.00	123.57	0.96
SO2-LC	5.00	6.00	1.00	7.00	10.00	0.98	11.00	11.14	0.96	37.00	61.43	0.98
SO2-MC	11.00	16.00	0.99	10.00	15.00	0.98	5.00	5.00	0.98	56.00	134.75	0.96

SO3-HC	4.00	4.00	1.00	6.00	6.00	1.00	6.00	9.00	0.94	59.00	97.15	0.97
SO3-LC	9.00	12.00	1.00	9.00	10.00	0.99	7.00	7.00	1.00	68.00	125.00	0.96
SO3-MC	8.00	8.33	1.00	9.00	10.00	0.99	8.00	11.00	0.91	49.00	82.33	0.98

Table S3.6. AMF virtual taxa of the 18S rRNA gene dataset before removing the singletons. OTUs sequences were compared in MaarjAM database.
This supporting table is available as Table S6 in the website of Environmental Microbiology Journal (http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13438/full).

Table S3.7. Kruskal–Wallis test on: (A) OTUs of AMF 18S rRNA gene, (B) OTUs of AMF ITS, (C) most abundant 20 OTUs of fungal ITS, (D) most abundant 50 OTUs of bacterial 16S rRNA gene.

A) Kruskal-Wallis test on AMF 18S rRNA gene taxa

	OTUs affiliation	Class level	Contamination level	plant species
OTU 1	*Diversispora sp. VTX00061*	Glomeromycetes	0.001905	0.9982
OTU 2	*Acaulospora acau10 VTX00028*	Glomeromycetes	0.1673	0.4342
OTU 3	*Glomus* sp.*VTX00265*	Glomeromycetes	0.8493	0.5176
OTU 4	*Glomus intraradices VTX00113*	Glomeromycetes	0.003496	0.6071
OTU 5	*Claroideoglomus* sp. *VTX00193*	Glomeromycetes	0.02143	0.6927

B) Kruskal-Wallis test on AMF ITS taxa

	OTUs affiliation	Class level	Contamination level	plant species
OTU 1	*Unclassified Diversispora*	Glomeromycetes	0.004537	0.6885
OTU 2	*Unclassified Rhizophagus*	Glomeromycetes	0.000751	0.925
OTU 3	*Glomus sp*	Glomeromycetes	0.4079	0.5636
OTU 4	*Glomus versiforme*	Glomeromycetes	0.03603	0.8301
OTU 5	*Entrophospora infrequens*	Glomeromycetes	0.1209	0.9274
OTU 6	*Unclassified Rhizophagus*	Glomeromycetes	0.0005533	0.7686
OTU 7	*Glomus sp*	Glomeromycetes	0.2739	0.7586
OTU 8	*Unclassified Acaulospora*	Glomeromycetes	0.1886	0.2502
OTU 9	*Glomus irregulare*	Glomeromycetes	0.03296	0.4602
OTU 10	*Claroideoglomus luteum*	Glomeromycetes	0.0005259	0.4422

C) Kruskal-Wallis test on the other ITS fungi taxa

	OTUs affiliation	Class level	Contamination level	plant species
OTU 1	*Pulvinula constellatio*	Pezizomycetes	0.008627	0.5001
OTU 2	*unclassified Ascomycota*	-	0.02184	0.3281

OTU 3	*Septoria citricola*	Dothideomycetes	0.1957	0.5233
OTU 4	*Spizellomyces plurigibbosus*	Chytridiomycetes	0.01543	0.7582
OTU 5	*Rhodotorula sp*	Urediniomycetes	0.1954	0.205
OTU 6	*unidentified Ascomycota*	-	0.08417	0.09763
OTU 7	*Fusarium sacchari*	Sordariomycetes	0.6838	0.1932
OTU 8	*unidentified*	-	0.1607	0.01797
OTU 9	*unclassified*	-	0.0154	0.481
OTU 10	*unclassified Ascomycota*	-	0.07152	0.2816
OTU 11	*unclassified Ascomycota*	-	0.7954	0.2689
OTU 12	*unclassified*	-	0.3679	0.3679
OTU 13	*Leptosphaeria sp*	Dothideomycetes	0.1581	0.09583
OTU 14	*unclassified*	-	0.3679	0.3679
OTU 15	*unclassified Pleosporales*	Dothideomycetes	0.5937	0.1253
OTU 16	*unclassified Basidiomycota*	-	0.1253	0.5937
OTU 17	*Geopora sp*	Pezizomycetes	0.3679	0.3679
OTU 18	*Rhodotorula glutinis*	Urediniomycetes	0.2398	0.7456
OTU 19	*unidentified*	-	0.3086	0.7195
OTU 20	*unclassified Sordariomycetes*	Sordariomycetes	0.1253	0.1253

D) Kruskal-Wallis test on the bacterial 16S rRNA gene taxa

	OTUs affeliation	**Class level**	**Contamination level**	**plant species**
OTU 1	*Pseudomonas*	Gammaproteobacteria	0.05137	0.3107
OTU 2	*Duganella*	Betaproteobacteria	0.1477	0.2048
OTU 3	*Pseudomonas*	Gammaproteobacteria	0.3751	0.529
OTU 4	*Pseudomonas*	Gammaproteobacteria	0.01856	0.03218
OTU 5	*Janthinobacterium*	Betaproteobacteria	0.1906	0.4143
OTU 6	*Enterobacter*	Gammaproteobacteria	0.02186	0.72
OTU 7	*Yersinia*	Gammaproteobacteria	0.6775	0.06549
OTU 8	*Streptomyces*	Actinobacteria	0.05386	0.07186
OTU 9	*Acidovorax*	Betaproteobacteria	0.7112	0.002563
OTU 10	*Streptomyces*	Actinobacteria	0.0425	0.1305
OTU 11	*Caulobacter*	Alphaproteobacteria	0.007955	0.4104
OTU 12	*Massilia*	Betaproteobacteria	0.9873	0.5411
OTU 13	*unclassified Burkholderiales*	Betaproteobacteria	0.1154	0.1494
OTU 14	*Massilia*	Betaproteobacteria	0.1124	0.206
OTU 15	*Lentzea*	Actinobacteria	0.8954	0.07729
OTU 16	*unclassified*	-	0.05453	0.8802
OTU 17	*unclassified Myxococcales*	Deltaproteobacteria	0.03708	0.276

OTU 18	*unclassified Myxococcales*	Deltaproteobacteria	0.1033	0.1075
OTU 19	*Roseateles*	Betaproteobacteria	0.2992	0.4269
OTU 20	*Flavobacterium*	Flavobacteria	0.05416	0.3444
OTU 21	*Streptomyces*	Actinobacteria	0.0264	0.02176
OTU 22	*Streptomyces*	Actinobacteria	0.02258	0.04187
OTU 23	*Lechevalieria*	Actinobacteria	0.2945	0.004914
OTU 24	*unclassified*	-	0.7751	0.04076
OTU 25	*Arthrobacter*	Actinobacteria	0.2672	0.07443
OTU 26	*Streptomyces*	Actinobacteria	0.03832	0.5952
OTU 27	*Bradyrhizobium*	Alphaproteobacteria	0.05465	0.4536
OTU 28	*Herbaspirillum*	Betaproteobacteria	0.02521	0.2302
OTU 29	*unclassified Rhodospirillales*	Alphaproteobacteria	0.5893	0.003573
OTU 30	*Sphingomonas*	Alphaproteobacteria	0.04077	0.3151
OTU 31	*unclassified*	-	0.19	0.7138
OTU 32	*Variovorax*	Betaproteobacteria	0.3483	0.8183
OTU 33	*Streptomyces*	Actinobacteria	0.0907	0.06903
OTU 34	*Lechevalieria*	Actinobacteria	0.2777	0.03598
OTU 35	*Aeromonas*	Gammaproteobacteria	0.4461	0.09937
OTU 36	*Aeromonas*	Gammaproteobacteria	0.03638	0.7278
OTU 37	*unclassified Myxococcales*	Deltaproteobacteria	0.113	0.09157
OTU 38	*Stenotrophomonas*	Gammaproteobacteria	0.04076	0.7708
OTU 39	*unclassified Betaproteobacteria*	Betaproteobacteria	0.4066	0.2544
OTU 40	*Leifsonia*	Actinobacteria	0.1285	0.06888
OTU 41	*Sphingobium*	Alphaproteobacteria	0.001386	0.7392
OTU 42	*Streptomyces*	Actinobacteria	0.1216	0.3784
OTU 43	*Sphingomonas*	Alphaproteobacteria	0.3679	0.3679
OTU 44	*Salinibacterium*	Actinobacteria	0.06033	0.01554
OTU 45	*Caulobacter*	Alphaproteobacteria	0.9072	0.03504
OTU 46	*Ideonella*	Betaproteobacteria	0.1093	0.01734
OTU 47	*Comamonas*	Betaproteobacteria	0.3578	0.03997
OTU 48	*Skermanella*	Alphaproteobacteria	0.04425	0.3641
OTU 49	*Pseudoxanthomonas*	Gammaproteobacteria	0.1888	0.2791
OTU 50	*unclassified Chromatiales*	Gammaproteobacteria	0.002228	0.1231

Table S3.8. *P*-values and adjusted *P*-values (*P*-val adj) of Spearman correlations test calculated between AMF genera and other fungal classes. *P*-value corrections were performed using false discovery rate method.

	Glomus		***Diversispora***		***Claroideoglomus***		***Acaulospora***	
	***P*-val**	***P*-val adj**	***P*-val**	***P*-val adj**	***P*-val**	***P*-val adj**	***P*-val**	***P*-val adj**
Agaricomycetes	0.503	0.849	0.552	0.849	0.800	0.872	0.831	0.872
Chytridiomycetes	0.021	0.166	0.003	0.069	0.003	0.069	0.452	0.823
Dothideomycetes	0.005	0.069	0.143	0.358	0.710	0.871	0.065	0.26
Leotiomycetes	0.012	0.126	0.050	0.254	0.685	0.871	0.858	0.872
Microbotryomycetes	0.296	0.593	0.536	0.849	0.872	0.872	0.869	0.872
Pezizomycetes	0.1855	0.422	0.050	0.254	0.618	0.871	0.714	0.871
Sordariomycetes	0.0782	0.274	0.057	0.256	0.856	0.872	0.531	0.849
Tremellomycetes	0.0491	0.254	0.095	0.274	0.429	0.818	0.675	0.871
Unclassified	0.0889	0.274	0.117	0.314	0.193	0.422	0.862	0.872
Ustilaginomycetes	0.6201	0.871	0.718	0.871	0.200	0.422	0.095	0.274

Table S3.9. *P*-values and adjusted *P*-values (*P*-val adj) of Spearman correlations test calculated between AMF genera and bacterial classes. *P*-value corrections were performed using false discovery rate method.

	Glomus		*Diversispora*		*Claroideoglomus*		*Acaulospora*	
	***P*-val**	***P*-val adj**	***P*-val**	***P*-val adj**	***P*-val**	***P*-val adj**	***P*-val**	***P*-val adj**
Gammaproteobacteria	0.3002	0.715	0.1953	0.705	0.6735	0.933	0.869	0.959
Betaproteobacteria	0.0565	0.66	0.8441	0.959	0.2548	0.705	0.1081	0.66
Actinobacteria	0.1214	0.66	0.2452	0.705	0.9626	0.98	0.7197	0.947
Alphaproteobacteria	0.1667	0.695	0.59	0.888	0.2229	0.705	0.3589	0.764
unclassified	0.9475	0.98	0.6743	0.933	0.5822	0.888	0.8557	0.959
Deltaproteobacteria	0.2688	0.705	0.097	0.66	0.4163	0.833	0.822	0.959
Acidobacteria_Gps	0.5701	0.888	0.8003	0.959	0.8161	0.959	0.4912	0.859
Flavobacteriia	0.9526	0.98	0.4357	0.844	0.5886	0.888	0.2541	0.705
Bacteroidetes_incertae_sedis	0.908	0.976	0.6117	0.888	0.8917	0.969	0.8731	0.959
Bacilli	0.0817	0.66	0.681	0.933	0.4509	0.851	0.016	0.534
Opitutae	0.8389	0.959	0.2054	0.705	0.1527	0.695	0.118	0.66
Sphingobacteriia	0.0804	0.66	0.2848	0.705	0.2891	0.705	0.1096	0.66
Gemmatimonadetes	0.5303	0.871	0.7656	0.959	0.4804	0.859	0.2309	0.705
Clostridia	0.3414	0.759	0.439	0.844	0.3715	0.773	0.9249	0.98
Anaerolineae	0.7185	0.947	0.2766	0.705	0.3554	0.764	0.6907	0.933
Cytophagia	0.1574	0.695	0.0035	0.332	0.0066	0.332	0.2703	0.705
Thermomicrobia	0.0779	0.66	0.2094	0.705	0.4607	0.853	0.5811	0.888
Subdivision3	0.0999	0.66	0.1135	0.66	0.3301	0.75	0.7858	0.959
Planctomycetia	0.6217	0.888	0.8407	0.959	0.0705	0.66	0.7313	0.95
Caldilineae	0.5989	0.888	0.498	0.859	0.9701	0.98	0.4979	0.859
Spartobacteria	0.1255	0.66	0.617	0.888	0.9503	0.98	0.5312	0.871
Chloroflexia	0.1428	0.695	0.1253	0.66	0.229	0.705	0.2553	0.705
Negativicutes	0.803	0.959	0.3134	0.729	0.0952	0.66	0.1253	0.66
Bacteroidia	0.803	0.959	0.5313	0.871	0.1615	0.695	0.2048	0.705
Candidatus_Hydrogenedens	1	1	0.8029	0.959	0.229	0.705	0.3788	0.773

Table S3.10. Concentrations of polycyclic aromatic hydrocarbons (PAH) and alkanes (C10-C50) measured from the rhizospheric soils of each plant species and contaminated basins. PAH and alkanes (C10-C50) were measured three times for each sample, then means and standard deviation were calculated for each sample. The data showed that basin 1 is the most contaminated site followed by basin 3, while basin 2 is the least contaminated site. Abbreviations mean: S-B: *Solidago canadensis*-Basin; P-B: *Populus balsamifera*-Basin; LB: *Lycopus europaeus.*-Basin; [C]: Concentrations; NA: Not detected.

PHCs contaminants	PHCs concentration in basin 1 (mg/kg of soil)				PHCs concentration in basin 2 (mg/kg of soil)				PHCs concentration in basin 3 (mg/kg of soil)			
	S-B-1	P-B-1	L-B-1	mean [C] in B-1	S-B-2	P-B-2	L-B-2	mean [C] in B-2	S-B-3	P-B-3	L-B-3	mean [C] in B-3
Anthracene	23	16	12	17±5.56	1.4	1.3	0.8	1.166±0.32	8.2	5	11	8.066±3
Benzo(a)anthracene	1.1	0.7	0.9	0.9±0.2	<0.1	<0.1	<0.1	NA	0.3	0.2	0.3	0.266±0.05
Benzo(a)pyrene	0.9	0.8	0.6	0.76±0.15	<0.1	<0.1	<0.1	NA	0.3	0.2	0.2	0.233±0.05
Benzo(b+j+k)fluoranthene	1.3	0.8	0.7	0.93±0.32	<0.1	<0.1	<0.1	NA	0.3	<0.1	0.2	NA
Benzo(ghi)perylene	0.5	0.4	0.3	0.4±0.1	<0.1	<0.1	<0.1	NA	0.2	0.1	0.2	0.166±0.05
Chrysene	2.3	1.3	1.6	1.73±0.51	<0.1	0.1	<0.1	NA	0.3	0.4	0.4	0.366±0.05
Fluoranthene	1.5	1.1	1.3	1.3±0.2	<0.1	<0.1	<0.1	NA	0.2	0.1	0.1	0.133±0.05
Fluorene	1.1	0.9	0.7	0.9±0.2	<0.1	<0.1	<0.1	NA	0.3	0.2	0.3	0.266±0.05
Phenanthrene	0.7	0.6	0.2	0.5±0.26	<0.1	<0.1	<0.1	NA	0.2	0.2	0.3	0.233±0.05
Pyrene	5.8	3.8	6.6	5.4±1.44	<0.1	0.2	<0.1	NA	1	0.7	0.4	0.7±0.3
Alkanes (C10-C50)	3000	1500	2100	2200±755	110	140	210	153.33±51.31	890	530	980	800±238.11

Table S3.11. OTUs of AMF ITS dataset before sub-sampling.
This supporting table is available as Table S11 in the website of Environmental Microbiology Journal (http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13438/full).

Table S3.12. OTUs of non AMF fungi dataset before sub-sampling.
This supporting table is available as Table S12 in the website of Environmental Microbiology Journal (http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13438/full).

Table S3.13. OTUs of bacteria 16S rRNA genes dataset before sub-sampling.
This supporting table is available as Table S13 in the website of Environmental Microbiology Journal (http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13438/full).

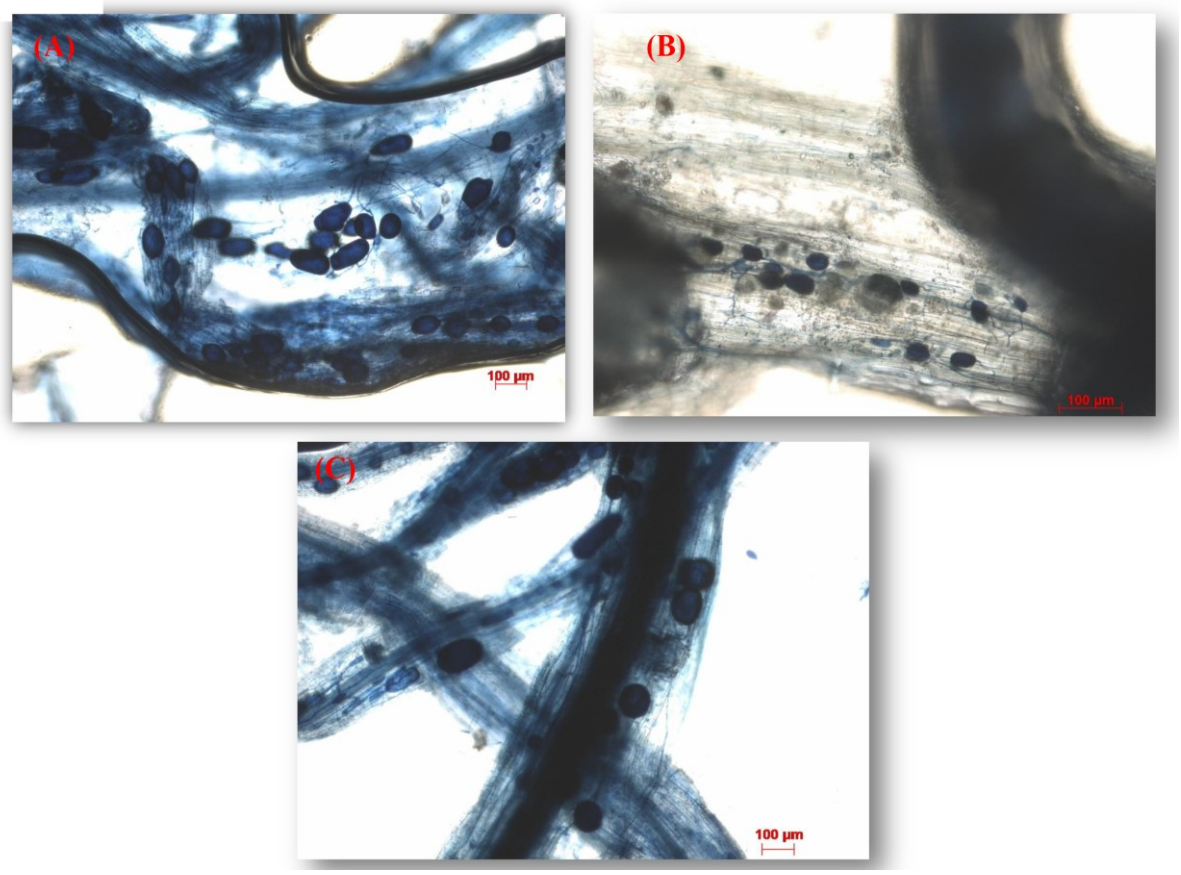


Figure S3.1. Root of *Solidago canadensis* (A), *Populus balsamifera* (B) and *Lycopus europaeus* (C) stained with trypan blue showing a high rate of mycorrhizal colonization.

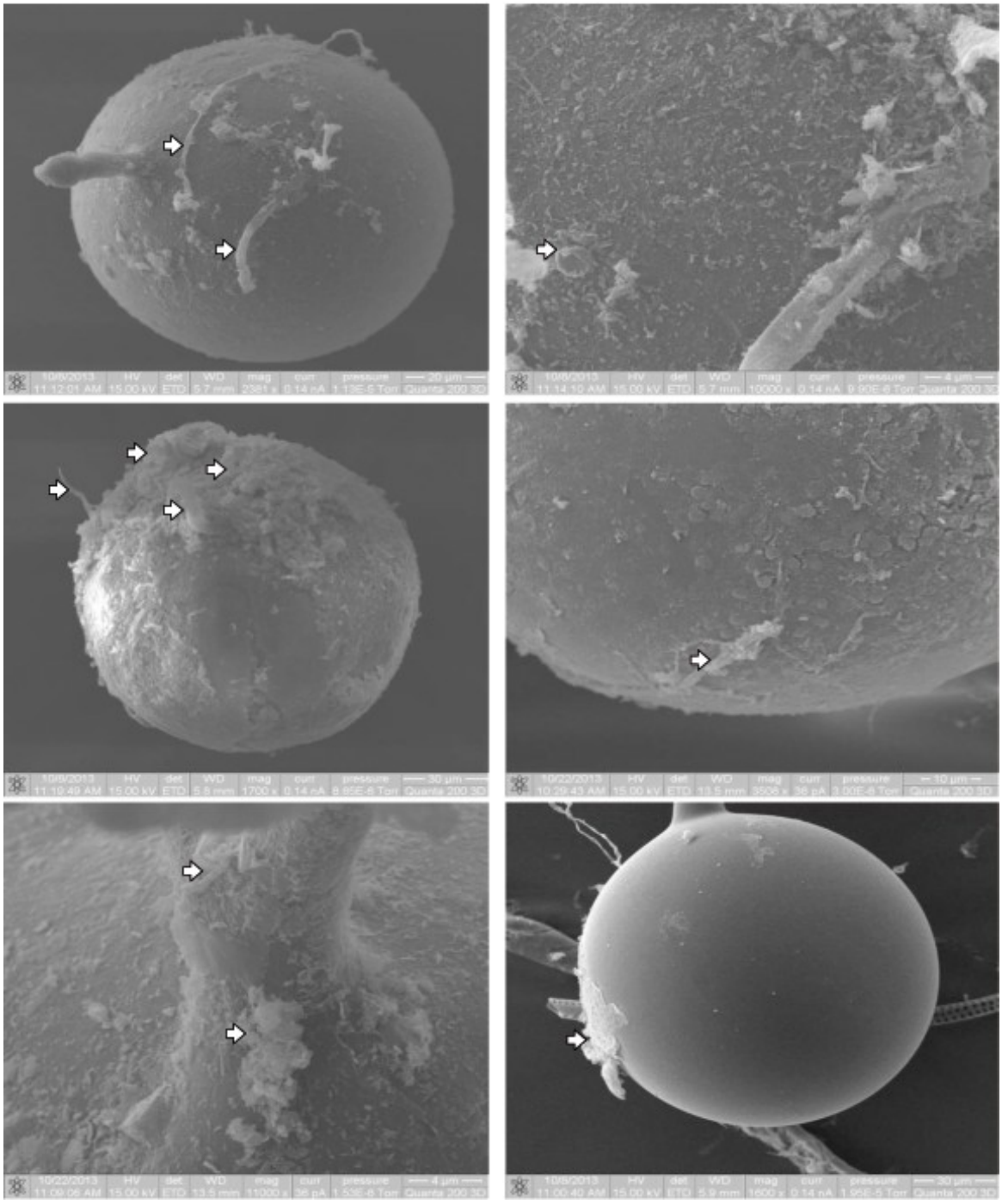


Figure S3.2. Scanning electron micrographs of AMF spores collected from the rhizospheric soils of the different plant species sampled from the contaminated basins. The white arrows showed biofilm-like structures of bacteria and mycelia of other fungi attached to the surface of AMF spores.

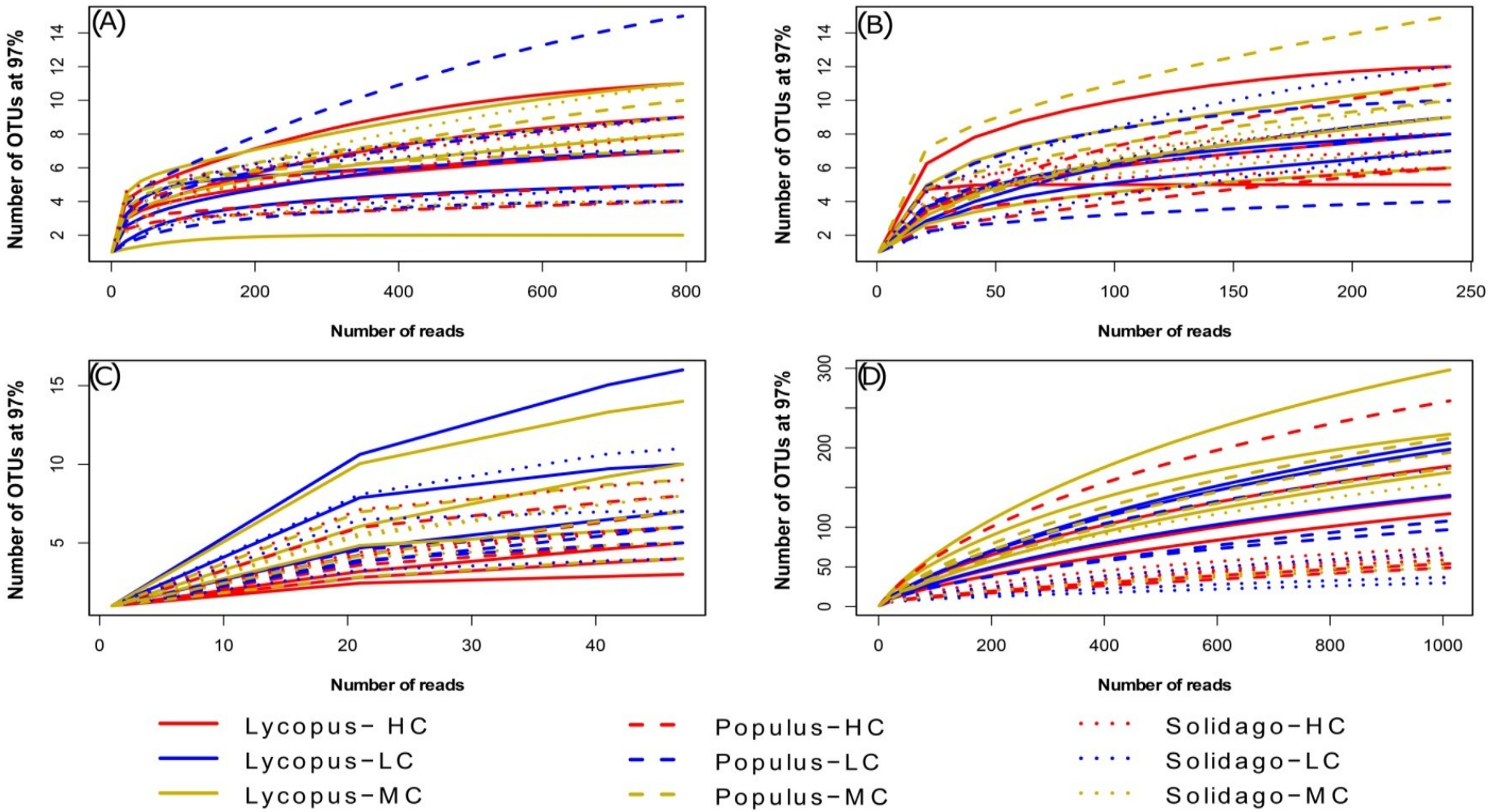


Figure S3.3. Rarefaction curve of OTUs for individual sample across the different datasets: (A) AMF 18S rRNA gene dataset, (B) AMF ITS dataset, (C) fungi ITS dataset (without AMF), (D) bacteria 16S rRNA gene dataset.

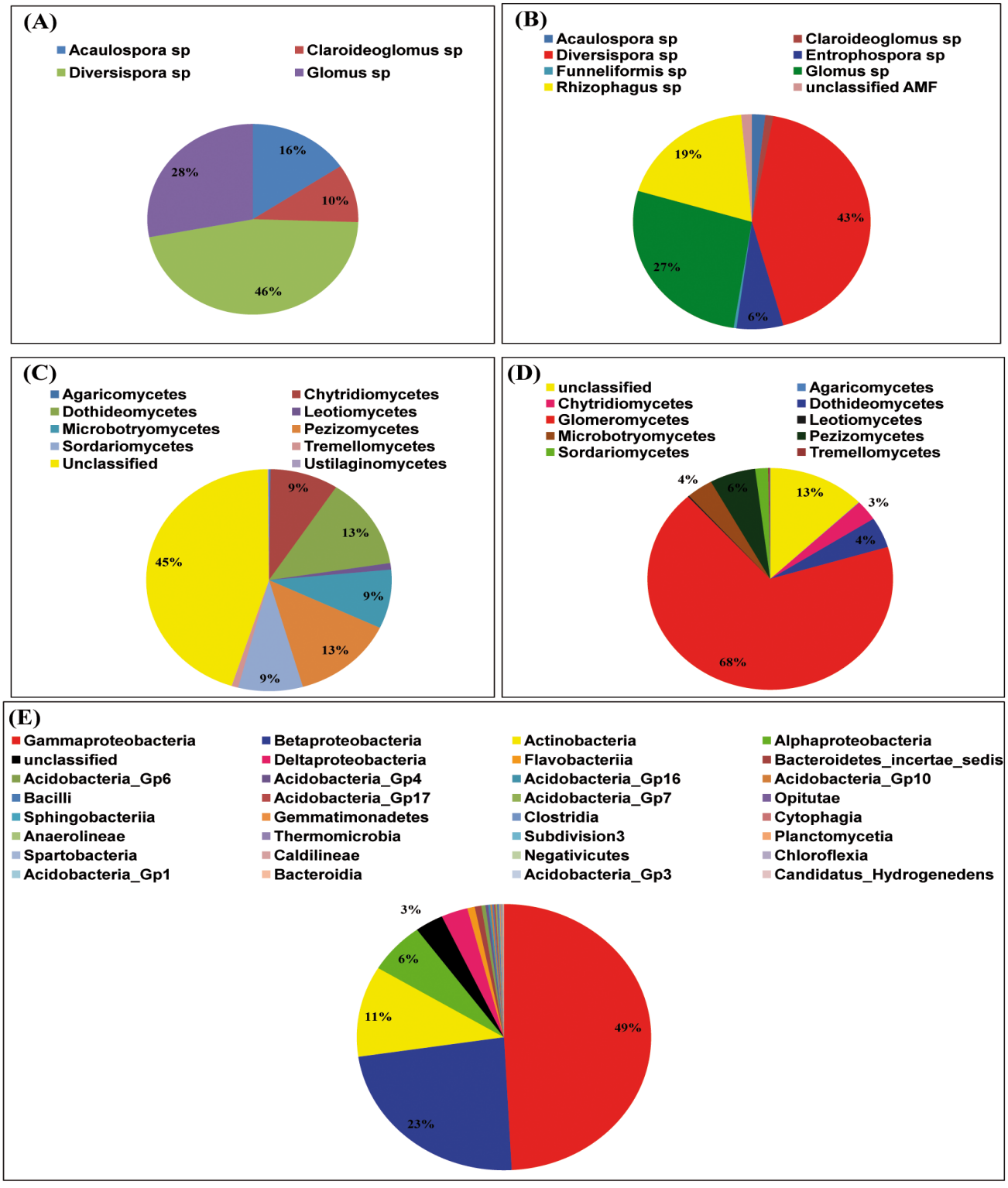


Figure S3.4. Proportion of the different: (A) AMF 18S rRNA gene genera, (B) AMF ITS genera, (C) non AM fungi classes, (D) total ITS fungi classes, and (E) bacterial 16S rRNA gene classes.

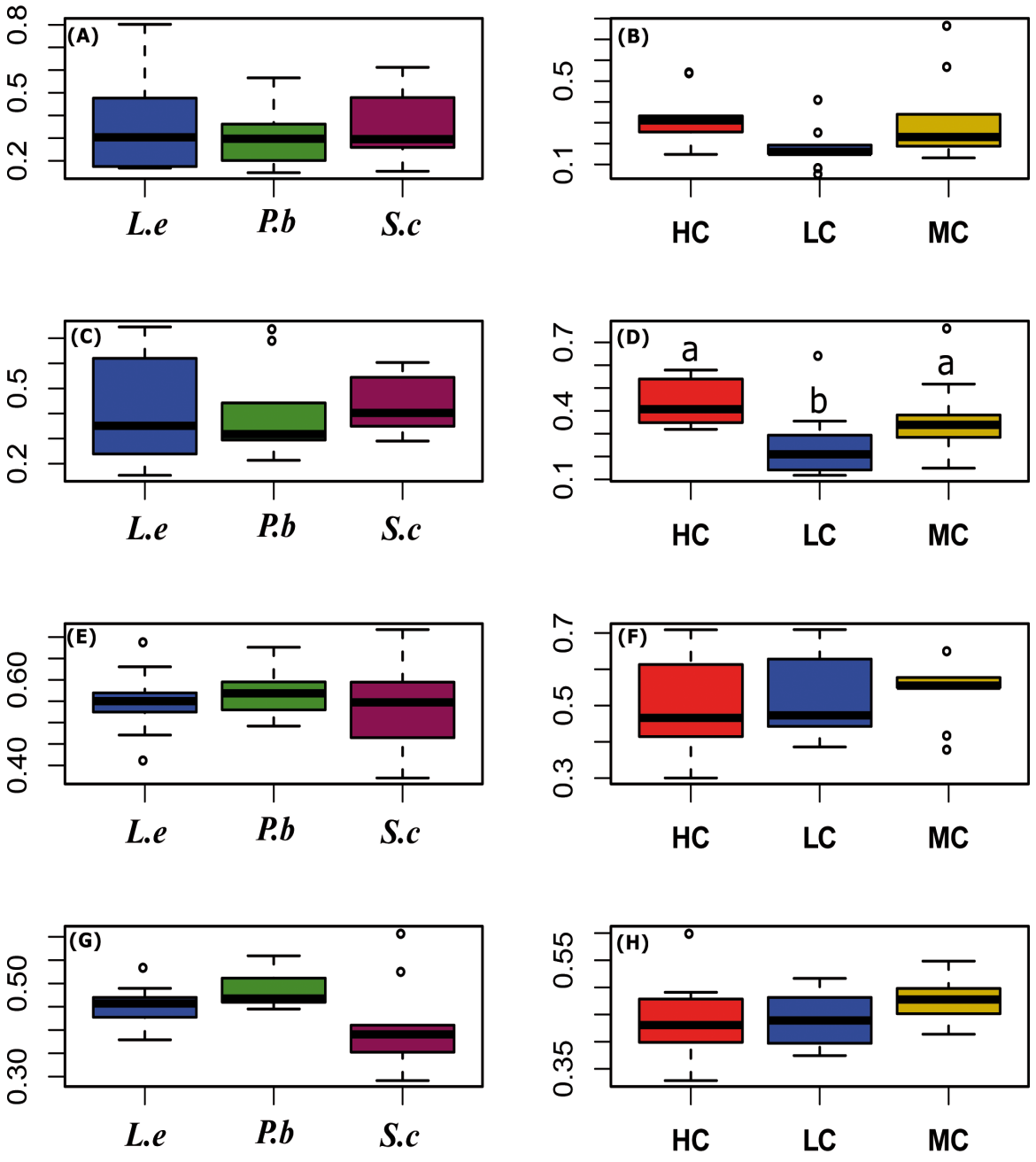


Figure S3.5. Boxplots of distance to centroid based on beta-dispersion analysis on the community structures of AMF, other fungi and bacteria against PHP concentration and plant species identity. The figures A, C, E and G showed the dispersion of AMF 18S rRNA gene, AMF ITS, other fungi ITS and bacteria 16S rRNA gene communities across plant species identity. The figures B, D, F and H showed the dispersion of AMF 18S rRNA gene, AMF ITS, other fungi ITS and bacteria 16S rRNA gene communities across PHP concentration. *S.c: S.canadensis*, *P.b*: *P.balsamifera*, *L.e*: *Lycopus europaeus*

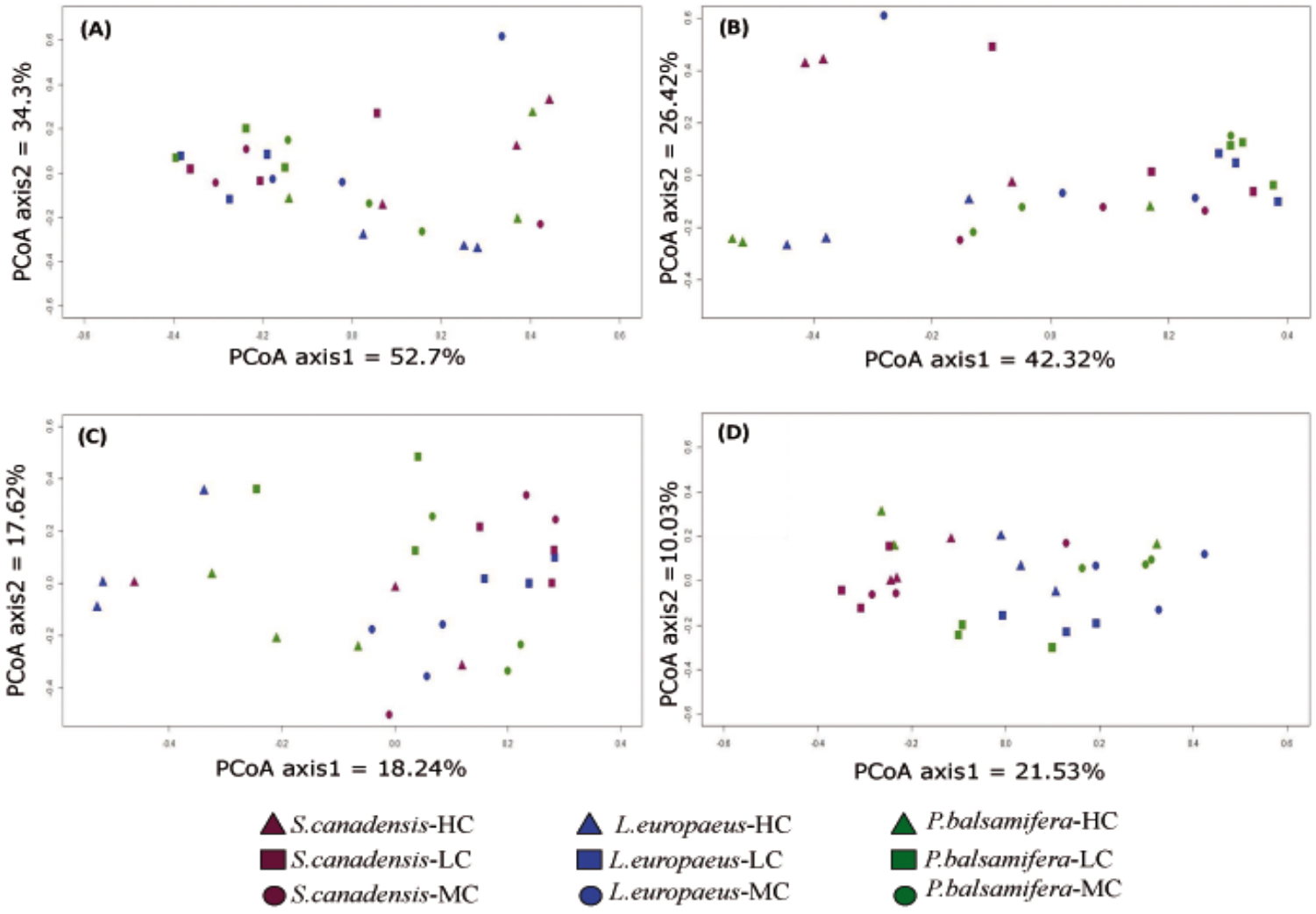


Figure S3.6. Principal coordinates analysis (PCoA) identical to that shown in Figure 2, but with color coding changed to show more clearly the community compositions assignments across plant species. (A) AMF 18S rRNA gene, (B) AMF ITS, (C) non AMF ITS, and (D) 16S rRNA gene.

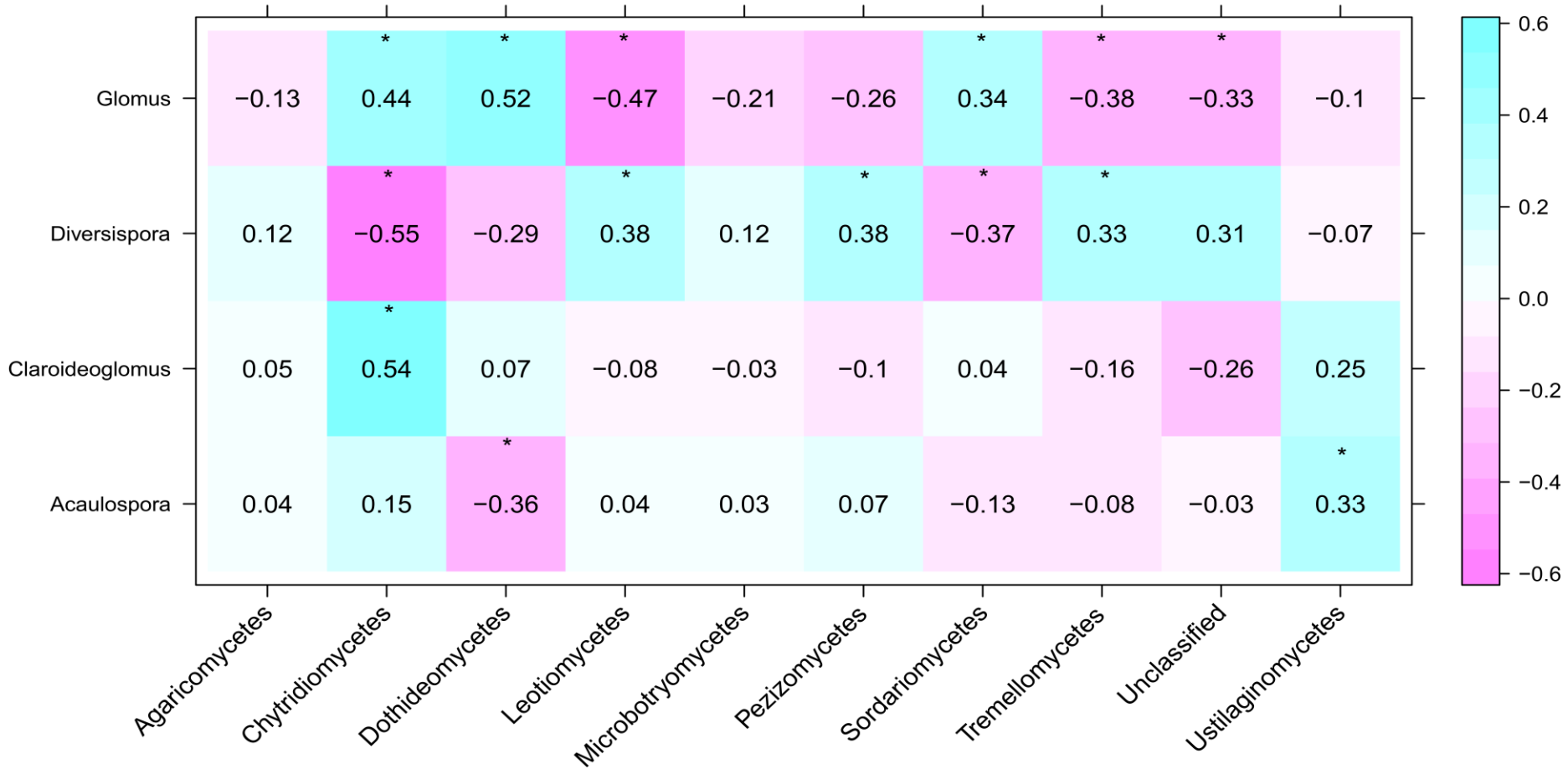


Figure S3.7. Color map of the Spearman's correlation coefficients between AMF genera and AMF-associated fungi classes. The green boxes indicate positive correlations, while the purple boxes indicate negative correlations. The black stars indicate that the *P*-value of the Spearman rank coefficients were lower than 0.1.

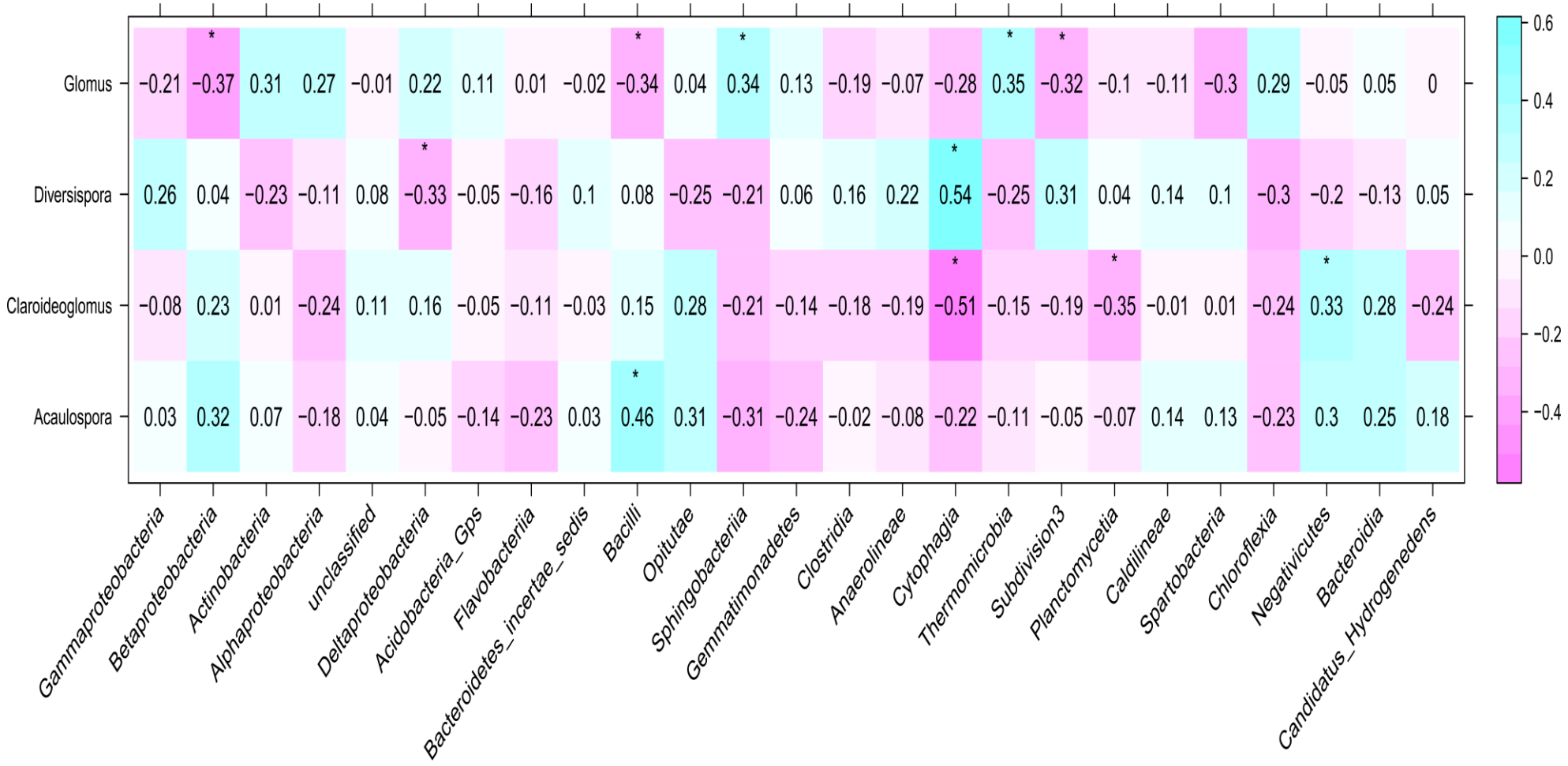


Figure S3.8. Color map of the Spearman's correlation coefficients between AMF genera and AMF-associated bacteria classes. The green boxes indicate positive correlations, while the purple boxes indicate negative correlations. The black stars indicate that the *P*-value of the Spearman rank coefficients were lower than 0.1.

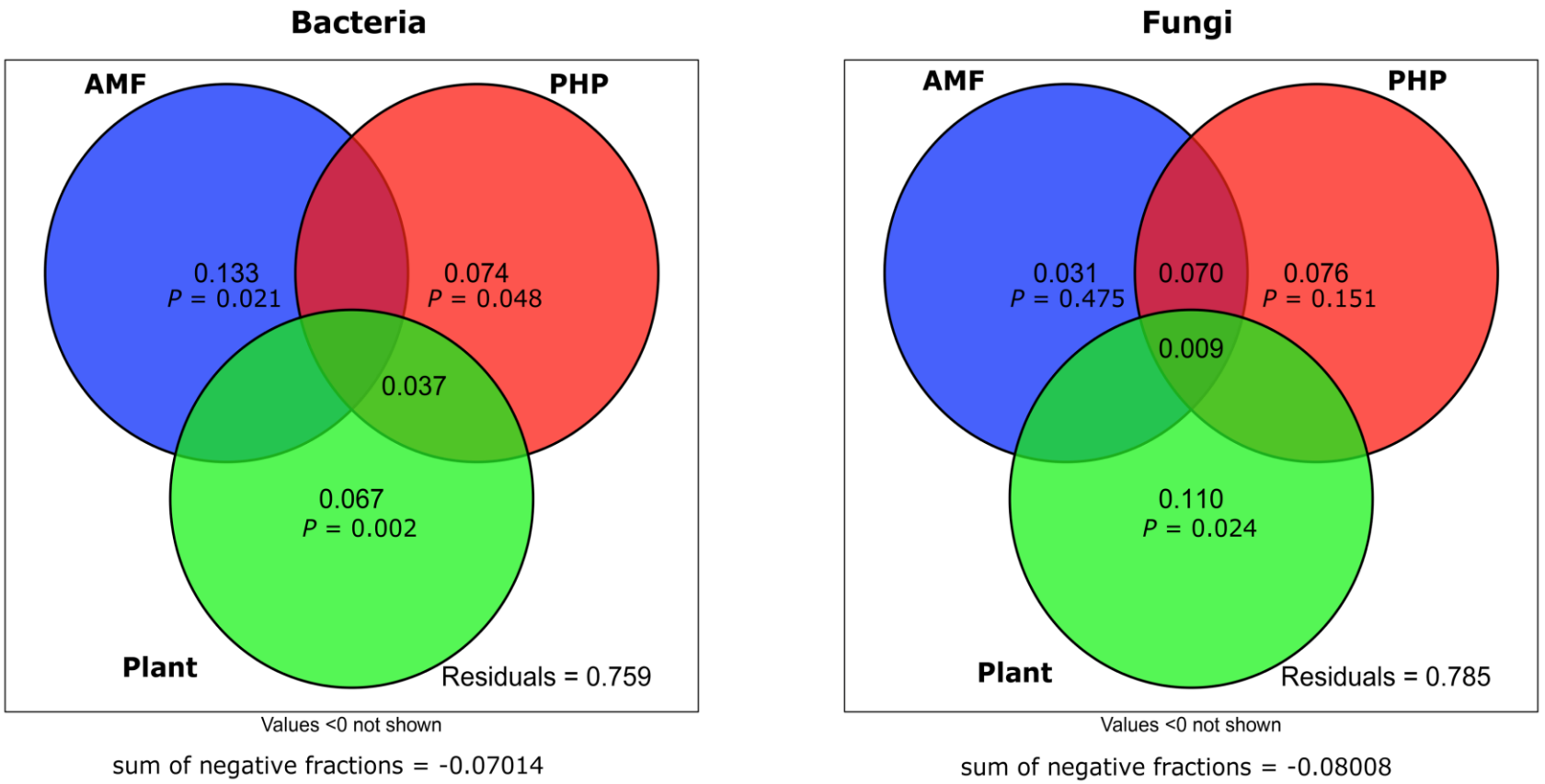


Figure S3.9. Variance partitioning of fungal and bacterial communities by AMF 18S rRNA gene, PHP concentration and plant species identity. The figures showed that the total variance explained by the three sets of explanatory matrices were 24.13% for the bacterial community variations (13.3% related AMF, 7.4% related to PHP concentrations and 6.7% related to plant species identity) and 21.51% for the fungal community variations (11% related to plant species identity, 7.6% related to PHP concentrations and 3.5% related to AMF).

3- Supplemental Material : Variance partitioning analysis (Chapter 3)

Because our results showed that the AMF spore-associated fungal and bacterial communities are linked to the AMF community structure, PHP concentration and plant species identity, a variance partitioning analysis was performed in order to look at the relative contribution of these three parameters (AMF community structure, PHP concentration and plant species identity) on the shifts in fungal and bacterial communities. The variance partitioning analysis was carried out using the "varpart" function in "vegan" package in R. Significance levels according to Monte Carlo permutation tests were carried out with 1000 permutations. The response variables used in this analysis were either the Hellinger-transformed matrix of fungal OTUs (represented by the OTUs whose sum of abundances > 50, totalizing 9 OTUs) or the Hellinger-transformed matrix of bacterial OTUs (represented by the OTUs whose sum of abundances > 100, totalizing 26 OTUs). The explanatory matrices were the Hellinger transformed matrix of AMF 18S rRNA gene (represented by the OTUs whose sum of abundances >50, totalizing 5 OTUs), the matrix of the three plant species (binary encoded) and the matrix of the PHP concentrations, which was represented by three columns (the first contained the concentrations of alkanes, the second the concentrations of anthracene, while the third contained the means concentration of the remaining PAH compounds reported in the Table S3.10.). The combination of remaining PAH compounds in a same column and the reduction of the number of columns (OTUs) were carried out in order to avoid collinearity problems caused by redundant variables, and in order to keep the number of rows (sites) higher than the number of columns (OTUs).

The results of this analysis showed that the three explanatory datasets explained a total contribution of 21.51 % of the variability in fungal communities and 24.13 % of the variability in bacterial communities. For the variation in bacterial communities, AMF matrix explained the highest fraction of the variation (13.3 %, P = 0.021), followed by PHP concentrations (7.4 %, P = 0.048) and plant species identity (6.7 %, P = 0.002). For the variation in fungi communities, plant species identity matrix was the only parameter that explained a significant contribution in fungal community variation (11.0 %, P = 0.024). PHP concentrations and AMF matrices explained only 7.6 % and 3.5 % in fungal community variation while their contribution was not significant (Figure S3.9.).

4- Supporting information (Chapter 4)

Variation in bacterial and fungal community structures across biotopes (soil vs roots), petroleum hydrocarbon concentration and plant species identity

Table S4.1. Observed richness, Chao1 estimator values and Good's coverage values for each individual sample across the different datasets. LY: *Lycopus europaeus*, PO: *Populus balsamifera*, SO: *Solidago Canadensis*

	Soil bacteria			Root bacteria			Soil fungi			Root fungi		
Samples	**sobs**	**chao**	**coverage**	**sobs**	**chao**	**coverage**	**sobs**	**chao**	**coverage**	**sobs**	**chao**	**coverage**
LY1-HC	357	648.20	0.749	67	161.23	0.438	73	90.00	0.981	43	48.14	0.982
LY1-LC	373	668.66	0.745	63	149.25	0.483	44	59.00	0.990	53	61.25	0.976
LY1-MC	370	669.16	0.740	65	137.07	0.472	59	85.00	0.986	65	76.40	0.963
LY2-HC	378	688.67	0.727	66	193.50	0.427	68	77.07	0.982	49	70.00	0.970
LY2-LC	442	770.81	0.688	71	204.00	0.360	44	53.00	0.990	46	53.09	0.974
LY2-MC	399	639.32	0.739	72	227.55	0.337	57	67.50	0.984	55	74.13	0.965
LY3-HC	398	854.79	0.691	76	284.00	0.270	64	75.77	0.981	43	56.20	0.976
LY3-LC	284	577.45	0.784	60	121.50	0.528	55	59.00	0.991	50	61.67	0.970
LY3-MC	385	736.65	0.713	73	220.50	0.326	48	54.60	0.987	58	71.60	0.967
PO1-HC	328	555.77	0.779	74	318.13	0.292	71	84.60	0.982	36	51.17	0.972
PO1-LC	433	746.75	0.706	75	299.00	0.281	43	51.25	0.987	39	42.11	0.984
PO1-MC	389	747.57	0.716	64	260.86	0.404	39	47.25	0.987	35	38.00	0.988
PO2-HC	330	589.21	0.777	77	361.75	0.236	39	59.00	0.983	38	68.33	0.972
PO2-LC	423	750.55	0.712	62	129.57	0.506	51	59.67	0.986	32	71.00	0.974
PO2-MC	378	795.24	0.709	67	279.14	0.382	53	66.20	0.987	36	40.00	0.984
PO3-HC	371	635.24	0.742	72	209.75	0.348	56	83.14	0.979	34	41.50	0.980
PO3-LC	397	673.72	0.730	68	140.06	0.438	55	68.13	0.984	33	39.43	0.980

PO3-MC	376	693.29	0.731	78	338.67	0.225	55	67.36	0.982	38	41.11	0.984
SO1-HC	299	566.05	0.788	55	185.00	0.551	74	92.75	0.974	39	44.08	0.976
SO1-LC	353	588.44	0.770	59	120.50	0.528	55	70.55	0.980	61	77.15	0.959
SO1-MC	301	479.89	0.812	54	103.58	0.607	61	66.00	0.989	49	73.43	0.963
SO2-HC	325	502.00	0.792	41	70.55	0.708	75	93.47	0.972	44	63.43	0.967
SO2-LC	422	799.35	0.691	47	81.36	0.685	64	81.00	0.981	50	63.13	0.970
SO2-MC	323	472.65	0.813	48	82.36	0.685	37	50.75	0.989	37	42.14	0.982
SO3-HC	343	558.18	0.773	50	142.63	0.562	66	83.50	0.978	54	112.00	0.943
SO3-LC	430	728.36	0.705	47	122.43	0.629	41	56.00	0.990	68	97.25	0.947
SO3-MC	333	576.79	0.768	50	112.33	0.618	43	52.43	0.987	28	31.00	0.986

Table S4.2. *P* values of Kruskal–Wallis test on the most abundant thirty OTUs of the: (A) 16S soil bacteria, (B) 16S root bacteria (C) ITS soil fungi, (D) ITS root fungi.

	OTUs affiliation	Class level	contamination level	plant species
A- Kruskal-Wallis test on the soil bacteria				
OTU 1 (C)	*Sphingomonas*	*Alphaproteobacteria*	0.0033	0.7989
OTU 2 (C)	*Unclassified Betaproteobacteria*	*Betaproteobacteria*	0.0003	0.9581
OTU 3 (P)	*Unclassified Rhizobiales*	*Alphaproteobacteria*	0.4508	0.0859
OTU 4	*Sphingomonas*	*Alphaproteobacteria*	0.3299	0.7332
OTU 5 (C)	*Unclassified Gammaproteobacteria*	*Gammaproteobacteria*	0.0399	0.614
OTU 6 (C)	*Skermanella*	*Alphaproteobacteria*	0.0138	0.7258
OTU 7	*Bradyrhizobium*	*Alphaproteobacteria*	0.4850	0.4854
OTU 8 (C)	*Caenimonas*	*Betaproteobacteria*	0.0678	0.5422
OTU 9	*Acinetobacter*	*Gammaproteobacteria*	0.2001	0.2091
OTU 10	*Bellilinea*	*Anaerolineae*	0.1566	0.1566
OTU 11 (C.P)	*Dongia*	*Alphaproteobacteria*	0.0582	0.0723
OTU 12 (C)	*Unclassified Acidobacteria Gp4*	*Acidobacteria_Gp4*	0.0002	0.8235
OTU 13 (C)	*Dongia*	*Alphaproteobacteria*	0.0567	0.4271
OTU 14	*Unclassified Deltaproteobacteria*	*Deltaproteobacteria*	0.8115	0.5443
OTU 15 (C)	*Steroidobacter*	*Gammaproteobacteria*	0.0051	0.9505
OTU 16	*Acinetobacter*	*Gammaproteobacteria*	0.3679	0.3678
OTU 17 (C)	*Unclassified Acidobacteria Gp4*	*Acidobacteria_Gp4*	0.0015	0.9458
OTU 18 (C)	*Unclassified Burkholderiales*	*Betaproteobacteria*	0.0555	0.6418
OTU 19 (C)	*Thermomonas*	*Gammaproteobacteria*	0.0067	0.4271
OTU 20 (C)	*Ferrovum*	*Betaproteobacteria*	0.0011	0.2729
OTU 21 (C)	*Unclassified Comamonadaceae*	*Betaproteobacteria*	0.0901	0.2268
OTU 22 (C)	*Unclassified Xanthomonadales*	*Gammaproteobacteria*	0.0273	0.7325
OTU 23 (C)	*Unclassified Burkholderiales*	*Betaproteobacteria*	0.0033	0.3724
OTU 24 (P)	*Unclassified Burkholderiales*	*Betaproteobacteria*	0.2148	0.0579
OTU 25	*Dongia*	*Alphaproteobacteria*	0.7484	0.3193
OTU 26	*Terrimonas*	*Sphingobacteriia*	0.5309	0.6835
OTU 27 (C)	*Unclassified Rhizobiales*	*Alphaproteobacteria*	0.0101	0.5183
OTU 28	*Skermanella*	*Alphaproteobacteria*	0.4622	0.3296
OTU 29 (C.P)	*Unclassified Ohtaekwangia*	*Bacteroidetes*	0.0264	0.0467
OTU 30 (C)	*Dongia*	*Alphaproteobacteria*	0.0135	0.5054
B- Kruskal-Wallis test on the root bacteria				
OTU 1(P)	*Pseudomonas*	*Gammaproteobacteria*	0.1154	0.0113
OTU 2 (C)	*Bradyrhizobium*	*Alphaproteobacteria*	0.0846	0.2454
OTU 3 (P)	*Pseudomonas*	*Gammaproteobacteria*	0.1115	0.0036

OTU 4 (C)	*Streptomyces*	*Actinobacteria*	0.0159	0.6044
OTU 5 (C)	*Steroidobacter*	*Gammaproteobacteria*	0.0049	0.8035
OTU 6 (C.P)	*Streptomyces*	*Actinobacteria*	0.0263	0.0602
OTU 7 (P)	*Actinoplanes*	*Actinobacteria*	0.2762	0.0581
OTU 8 (C.P)	*Sphingomonas*	*Alphaproteobacteria*	0.0737	0.0470
OTU 9 (P)	*Duganella*	*Betaproteobacteria*	0.2687	0.0005
OTU 10 (C)	*Rhizobacter*	*Gammaproteobacteria*	0.0416	0.2374
OTU 11 (P)	*Lentzea*	*Actinobacteria*	0.7349	0.0821
OTU 12	*Unclassified Burkholderiales*	*Betaproteobacteria*	0.1818	0.1390
OTU 13 (P)	*Unclassified Proteobacteria*	*Unclassified Proteobacteria*	0.4253	0.0018
OTU 14	*Leifsonia*	*Actinobacteria*	0.4210	0.9354
OTU 15 (P)	*Hyphomicrobium*	*Alphaproteobacteria*	0.6685	0.0535
OTU 16 (C)	*Skermanella*	*Alphaproteobacteria*	0.0939	0.2796
OTU 17 (P)	*Hyphomicrobium*	*Alphaproteobacteria*	0.8675	0.0667
OTU 18 (P)	*Steroidobacter*	*Gammaproteobacteria*	0.3218	0.0021
OTU 19 (C)	*Sphingobium*	*Alphaproteobacteria*	0.0932	0.3735
OTU 20 (P)	*Limnobacter*	*Betaproteobacteria*	0.4367	0.0097
OTU 21 (P)	*Pseudomonas*	*Gammaproteobacteria*	0.1985	0.0121
OTU 22 (C)	*Actinoplanes*	*Actinobacteria*	0.0540	0.4537
OTU 23 (P)	*Ideonella*	*Betaproteobacteria*	0.1494	0.0113
OTU 24 (P)	*Altererythrobacter*	*Alphaproteobacteria*	0.2924	0.0012
OTU 25	*Unclassified Gammaproteobacteria*	*Gammaproteobacteria*	0.9830	0.2018
OTU 26 (C)	*Hoeflea*	*Alphaproteobacteria*	0.0906	0.1533
OTU 27 (C)	*Actinoplanes*	*Actinobacteria*	0.0405	0.1917
OTU 28 (C)	*Ideonella*	*Betaproteobacteria*	0.0637	0.8015
OTU 29	*Acidobacteria Gp6*	*Acidobacteria_Gp6*	0.2433	0.1528
OTU 30	*Dongia*	*Alphaproteobacteria*	0.2673	0.2673
C- Kruskal-Wallis test on the soil fungi				
OTU 1	*Fusarium sacchari*	*Sordariomycetes*	0.1139	0.4173
OTU 2 (C.P)	*Unclassified Pleosporales*	*Dothideomycetes*	0.0912	0.0126
OTU 3 (P)	*Unclassified Thelephoraceae*	*Agaricomycetes*	0.1226	0.0160
OTU 4	*Unclassified Sordariomycetes*	*Sordariomycetes*	0.1315	0.7498
OTU 5 (C)	*Emericellopsis sp*	*Sordariomycetes*	0.0004	0.8375
OTU 6 (C)	*Penicillium sp*	*Eurotiomycetes*	0.0016	0.5295
OTU 7	*Cladosporium sp*	*Dothideomycetes*	0.4228	0.5873
OTU 8 (C)	*Spizellomyces plurigibbosus*	*Chytridiomycetes*	0.0002	0.9721
OTU 9 (C)	*Unclassified fungi*	*Unclassified fungi*	0.0090	0.5157
OTU 10	*Unclassified Sordariales*	*Sordariomycetes*	0.2995	0.8009
OTU 11 (C)	*Leptosphaeria sp*	*Dothideomycetes*	0.0546	0.1132

OTU 12 (C)	*endophytic ascomycete sp*	*Unidentified ascomycete*	0.0068	0.2526
OTU 13 (C)	*Unclassified Lasiosphaeriaceae*	*Sordariomycetes*	0.0003	0.9631
OTU 14 (P)	*Pleosporaceae sp*	*Dothideomycetes*	0.1683	0.0209
OTU 15 (C)	*Acremonium sp*	*Sordariomycetes*	0.0031	0.9489
OTU 16	*fungal sp QLF106*	*Unidentified fungi*	0.6794	0.1530
OTU 17	*Alternaria sp*	*Dothideomycetes*	0.7618	0.3890
OTU 18 (P)	*Unclassified Basidiomycota*	*Unclassified Basidiomycota*	0.5688	0.0034
OTU 19 (C)	*Fusarium sp*	*Sordariomycetes*	0.0545	0.6308
OTU 20	*Cadophora luteo olivacea*	*Leotiomycetes*	0.4874	0.6427
OTU 21 (C)	*Phoma herbarum*	*Dothideomycetes*	0.0847	0.6056
OTU 22	*Mortierella alpina*	*Incertae_sedis*	0.1377	0.3057
OTU 23	*Chalara sp*	*Incertae_sedis*	0.5635	0.1922
OTU 24	*Uncultured Ganoderma*	*Agaricomycetes*	0.5943	0.7733
OTU 25 (C)	*Sphaerosporella brunnea*	*Pezizomycetes*	0.0113	0.1154
OTU 26	*Cladosporium cladosporioides*	*Dothideomycetes*	0.1234	0.5300
OTU 27 (C)	*Pycnidiophora sp*	*Dothideomycetes*	0.0005	0.6691
OTU 28 (P)	*Dioszegia changbaiensis*	*Tremellomycetes*	0.7342	0.0885
OTU 29 (C)	*Unclassified fungi*	*Unclassified fungi*	0.0008	0.5646
OTU 30	*Unclassified fungi*	*Unclassified fungi*	0.3812	0.3257
	D- Kruskal-Wallis test on the root fungi			
OTU 1 (P)	*Leptosphaeria sp*	*Dothideomycetes*	0.1705	0.0006
OTU 2	*Unclassified Ascomycota*	*Unclassified Ascomycota*	0.9606	0.8872
OTU 3 (P)	*fungal sp QLF106*	*Unidentified fungi*	0.7043	0.0003
OTU 4 (C)	*Entrophospora infrequens*	*Glomeromycetes*	<0.0001	0.8253
OTU 5 (C.P)	*Unclassified Sebacinaceae*	*Agaricomycetes*	0.0099	0.0165
OTU 6 (P)	*Olpidium brassicae*	*Chytridiomycetes*	0.2788	<0.0001
OTU 7 (C.P)	*Phoma herbarum*	*Dothideomycetes*	0.0014	0.0224
OTU 8 (C)	*Unclassified Pleosporales*	*Dothideomycetes*	0.0405	0.6149
OTU 9 (C)	*Unclassified Helotiales*	*Leotiomycetes*	0.0003	0.4083
OTU 10 (C.P)	*Uncultured Claroideoglomus*	*Glomeromycetes*	0.0105	0.0272
OTU 11 (P)	*Uncultured Glomus*	*Glomeromycetes*	0.3518	0.0007
OTU 12 (C)	*fungal sp SUN1*	*Unidentified fungi*	0.0073	0.1138
OTU 13 (C)	*Unidentified Ascomycota*	*Unidentified Ascomycota*	0.0002	0.1600
OTU 14 (C.P)	*Chalara sp*	*Incertae_sedis*	0.0007	0.0730
OTU 15 (C.P)	*Pleosporales sp*	*Dothideomycetes*	0.0393	0.0393
OTU 16 (P)	*Rhizophagus irregularis*	*Glomeromycetes*	0.8907	0.0001
OTU 17 (P)	*Unclassified Glomeraceae*	*Glomeromycetes*	0.1677	0.0032
OTU 18 (C.P)	*Unclassified Glomeraceae*	*Glomeromycetes*	0.0584	0.0565

OTU 19 (P)	*Unclassified Thelephoraceae*	*Agaricomycetes*	0.8501	0.0012
OTU 20 (P)	*Cadophora luteo olivacea*	*Leotiomycetes*	0.2541	0.0006
OTU 21	*Alternaria sp*	*Dothideomycetes*	0.9722	0.2991
OTU 22 (C)	*Fusarium sacchari*	*Sordariomycetes*	0.0674	0.8795
OTU 23 (C)	*Pleosporales sp*	*Dothideomycetes*	0.0004	0.8035
OTU 24	*Glomus sp*	*Glomeromycetes*	0.1217	0.1217
OTU 25 (C.P)	*Uncultured Tetracladium*	*unidentified*	0.0036	0.0266
OTU 26 (C)	*Pulvinula constellatio*	*Pezizomycetes*	0.0019	0.1114
OTU 27 (P)	*Dioszegia changbaiensis*	*Tremellomycetes*	0.7417	<0.0001
OTU 28 (C.P)	*Spizellomyces plurigibbosus*	*Chytridiomycetes*	0.0453	0.0224
OTU 29 (C)	*Myrothecium sp*	*Sordariomycetes*	0.0020	0.2011
OTU 30 (P)	*Podospora communis*	*Sordariomycetes*	0.2025	0.0654

(C): significant effect across contamination level, (P): significant effect across plant species, (C.P): significant effect across contamination level and plant species

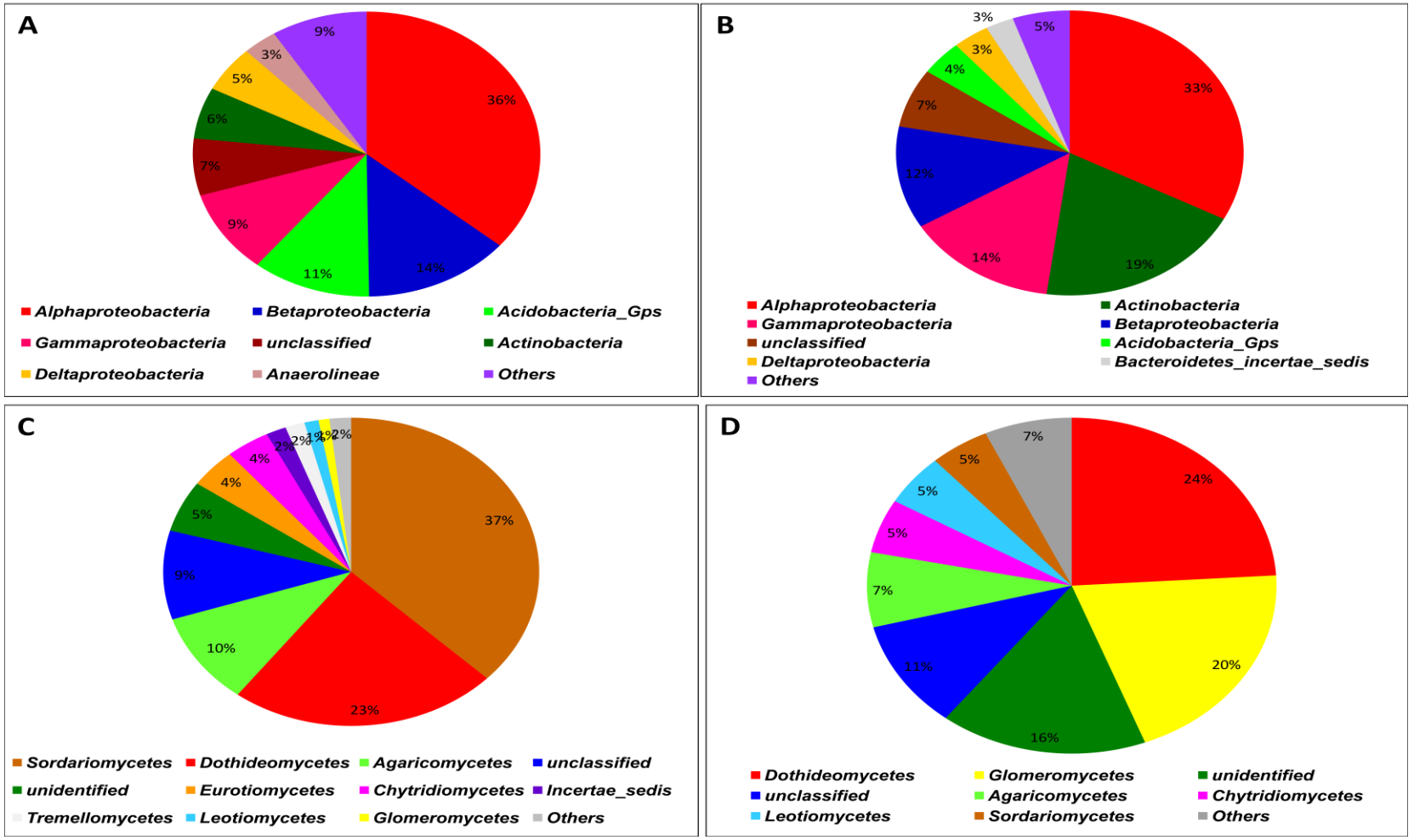


Figure S4.1. Proportion of the different: (A) soil bacteria classes, (B) root bacteria classes, (C) soil fungi classes, (D) root fungi classes.

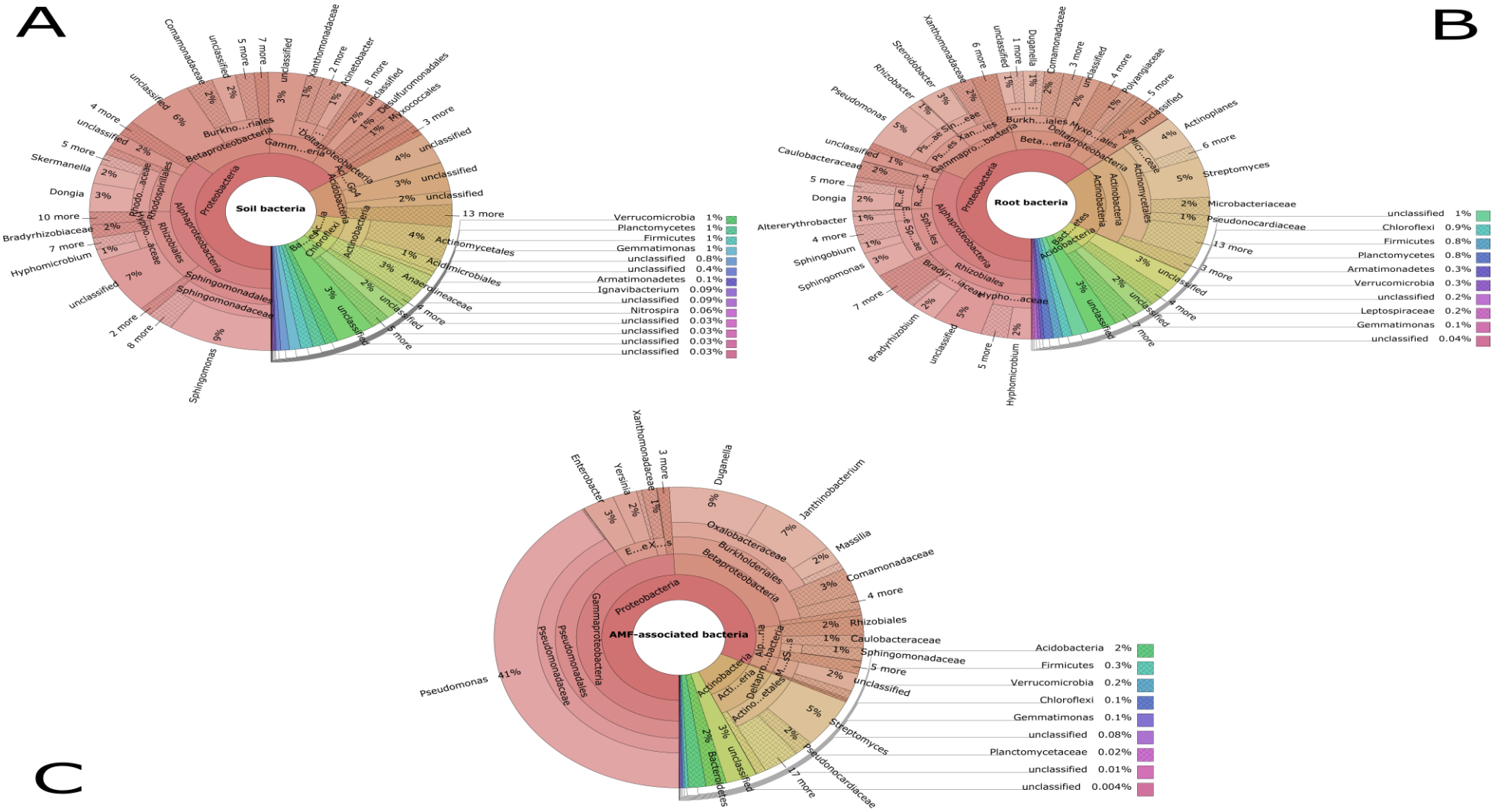


Figure S4.2. Krona charts showing the taxonomic identification and relative abundance of: (A) soil bacteria, (B) root bacteria and (C) AMF-associated bacteria.

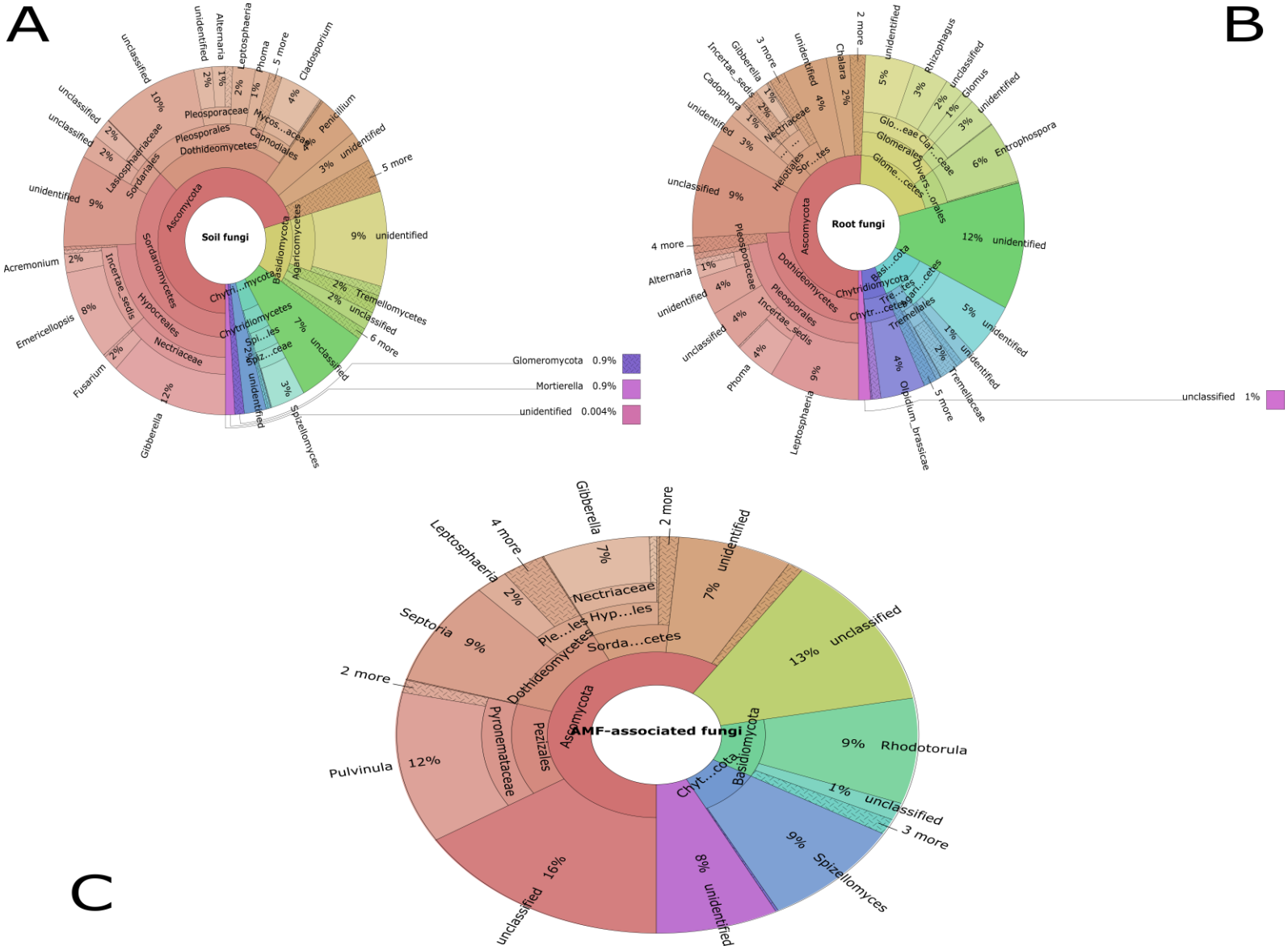


Figure S4.3. Krona charts showing the taxonomic identification and relative abundance of: (A) soil fungi, (B) root fungi and (C) AMF-associated fungi.

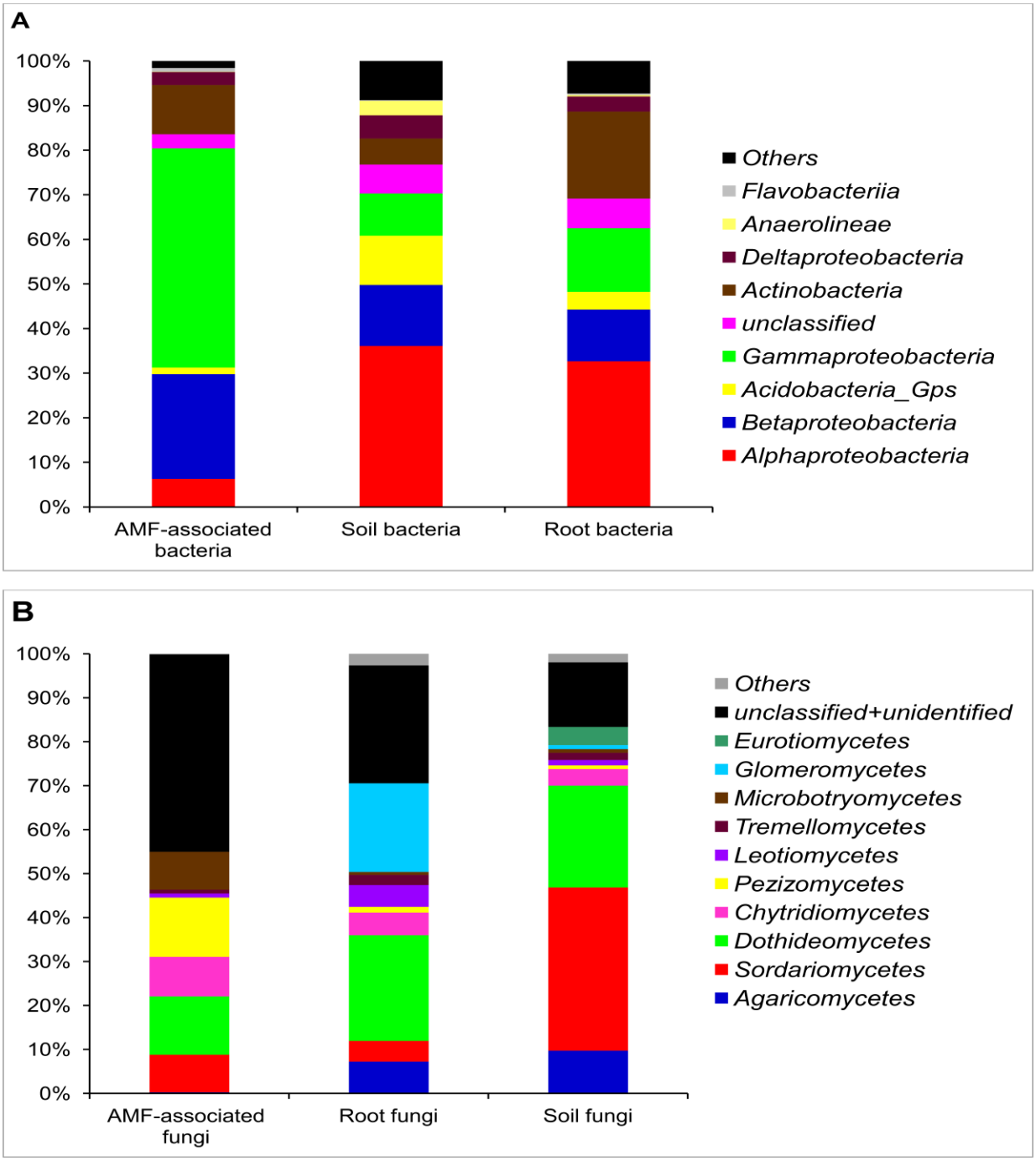


Figure S4.4. comparison of relative abundances between: (A) the bacteria classes found in soil, roots and in association with AMF spores; (B) the fungi classes found in soil, roots and in association with AMF spores.

5- Article 1 - Beaudet *et al.*, 2013

OPEN ACCESS Freely available online

PLOS ONE

Rapid Mitochondrial Genome Evolution through Invasion of Mobile Elements in Two Closely Related Species of Arbuscular Mycorrhizal Fungi

Denis Beaudet, Maryam Nadimi, Bachir Iffis, Mohamed Hijri*

Département de Sciences Biologiques, Institut de Recherche en Biologie Végétale, Université de Montréal, Montréal, Québec, Canada

Abstract

Arbuscular mycorrhizal fungi (AMF) are common and important plant symbionts. They have coenocytic hyphae and form multinucleated spores. The nuclear genome of AMF is polymorphic and its organization is not well understood, which makes the development of reliable molecular markers challenging. In stark contrast, their mitochondrial genome (mtDNA) is homogeneous. To assess the intra- and inter-specific mitochondrial variability in closely related *Glomus species*, we performed 454 sequencing on total genomic DNA of *Glomus sp.* isolate DAOM-229456 and we compared its mtDNA with two *G. irregulare* isolates. We found that the mtDNA of *Glomus sp.* is homogeneous, identical in gene order and, with respect to the sequences of coding regions, almost identical to *G. irregulare*. However, certain genomic regions vary substantially, due to insertions/deletions of elements such as introns, mitochondrial plasmid-like DNA polymerase genes and mobile open reading frames. We found no evidence of mitochondrial or cytoplasmic plasmids in *Glomus* species, and mobile ORFs in *Glomus* are responsible for the formation of four gene hybrids in *atp6*, *atp9*, *cox2*, and *nad3*, which are most probably the result of horizontal gene transfer and are expressed at the mRNA level. We found evidence for substantial sequence variation in defined regions of mtDNA, even among closely related isolates with otherwise identical coding gene sequences. This variation makes it possible to design reliable intra- and inter-specific markers.

Citation: Beaudet D, Nadimi M, Iffis B, Hijri M (2013) Rapid Mitochondrial Genome Evolution through Invasion of Mobile Elements in Two Closely Related Species of Arbuscular Mycorrhizal Fungi. PLoS ONE 8(4): e60768. doi:10.1371/journal.pone.0060768

Editor: Cecile Fairhead, Institut de Genetique et Microbiologie, France

Received December 6, 2012; **Accepted** March 2, 2013; **Published** April 18, 2013

Funding: This work is a part of a research project organized and coordinated by Premier Tech. The authors are grateful for financial support from NSERC Cooperative Research and Development (CRD) grant number RDPJ 395241-09, Premier Tech and CRIBIQ. The authors would also like to thank Biopterre centre développement des bioproduits and CRBM for their support. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors declare that they have received funds from Premier Tech (a profite company) in a form of a matching funds along with the Natural Sciences and Engineering Research Council of Canada. The authors are not employed nor consultants in Premier Tech and they are not patenting any of the results presented in this manuscript. The authors confirm that this does not alter their adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: Mohamed.Hijri@Umontreal.ca

Introduction

Arbuscular mycorrhizal fungi (AMF) are plant root-inhabiting obligate symbionts that form symbiotic associations with approximately 80% of plant species [1,2]. This symbiosis helps plants to acquire nutrients and protects them from soil-borne pathogens [3,4] by inducing plant resistance [5–8] or inhibiting pathogen growth [9]. In return, plants provide carbohydrates, which AMF cannot acquire from extracellular sources. They are an important component of soil microbial communities, as they are able to exchange their genetic material between compatible isolates through a process called anastomosis [10]. The latter have been hypothesized to be an important factor in maintaining the genetic diversity found in Glomeromycota and to attenuate the effect of genetic drift within a population [10–14]. AMF are currently thought to reproduce clonally, based on the absence of a recognizable sexual stage (or apparatus). However, this hypothesis has been challenged by the identification of many orthologues of sexually-related genes [15–17], which suggests at least the presence of cryptic recombination. AMF spores and hyphae are multinucleated, but their true genetic organization is currently under debate [11,18–21]. However, evidence strongly suggests that nuclei can be genetically divergent within an AMF individual. Thus, AMF are characterized by considerable within-isolate nuclear genetic diversity even at the expression level [22]. The presence of such diversity in AMF individuals/populations [22,23], combined with a lack of molecular data, have hindered the use of nuclear markers to assess questions on community structure, diversity and function. In contrast, AMF mitochondrial (mt) DNA is homogeneous within single isolates [24,25], making it a good target for marker development. Following this logic, the mitochondrial large subunit (LSU) rRNA gene has been explored for its usefulness as a marker [24,26,27], although determining its specificity at the isolate level is still challenging for all AMF taxa aside from the model species *G. irregulare*.

Comparative AMF mitochondrial genomics has been proposed as an approach to open up new possibilities for development of strain-specific molecular markers [25,28,29] given that the type of mitochondrial marker necessary to establish specificity at different divergence levels may vary. This approach has been shown to be a powerful tool for the study of evolutionary relationships among lower fungi [30]. Unfortunately, only three AMF mitochondrial genomes had been published until recently, including that of *Glomus intraradices* [25,29] (renamed to *G. irregulare* [31] and changed again recently to *Rhizophagus irregularis* based on an

exhaustive molecular phylogeny of rRNA genes [32]; in the present paper, we will use the older nomenclature) as well as those of two distant AMF species, *Gigaspora rosea* and *Gigaspora margarita* [33,34]. Compared to *G. irregulare*, the Gigasporaceae genomes have an inflated mitochondrial genome size that is mainly the result of extended intergenic regions. These regions are not syntenic and both genomes harbor *cox1* and *rns* genes with exons encoded on different strands, whose products are joined at the RNA level through either trans-splicing events of group I introns, or base-pairing. The mitochondrial protein sequences in the dataset were sufficient to confirm the phylogenetic relationship of AMF with *Mortierellales* as a sister group. This shows that a broader sampling of AMF mtDNA can answer questions about the evolution of these ecologically important fungi. Formey et al. (2012) have recently sequenced four isolates of *G. irregulare* [29] and were able to develop isolate-specific markers using variable regions That were created by the insertion of mobile elements.

Those elements, including linear or circular plasmids and mobile ORF encoding endonucleases (mORFs), are present in a broad range of fungal mitochondrial genomes (For review see [35]). Plasmids are autonomously-replicating circular or linear extrachromosomal DNA molecules. They are found in three broad types: circular plasmids encoding a DNA polymerase gene *(dpo)* [36], linear plasmids with terminal inverted repeats encoding either a *dpo* or *rpo* (RNA polymerase) gene or both [37], and retroplasmids, which usually encode a reverse transcriptase [38]. Free linear or circular plasmids encoding *dpo* can be present in the mitochondria of fungi [36] and plants [39]. Segments have been shown to integrate within the mtDNA of fungi [40–42], but plasmid-related *dpo* insertions tend to fragment, shorten (since they are not selected for) and eventually

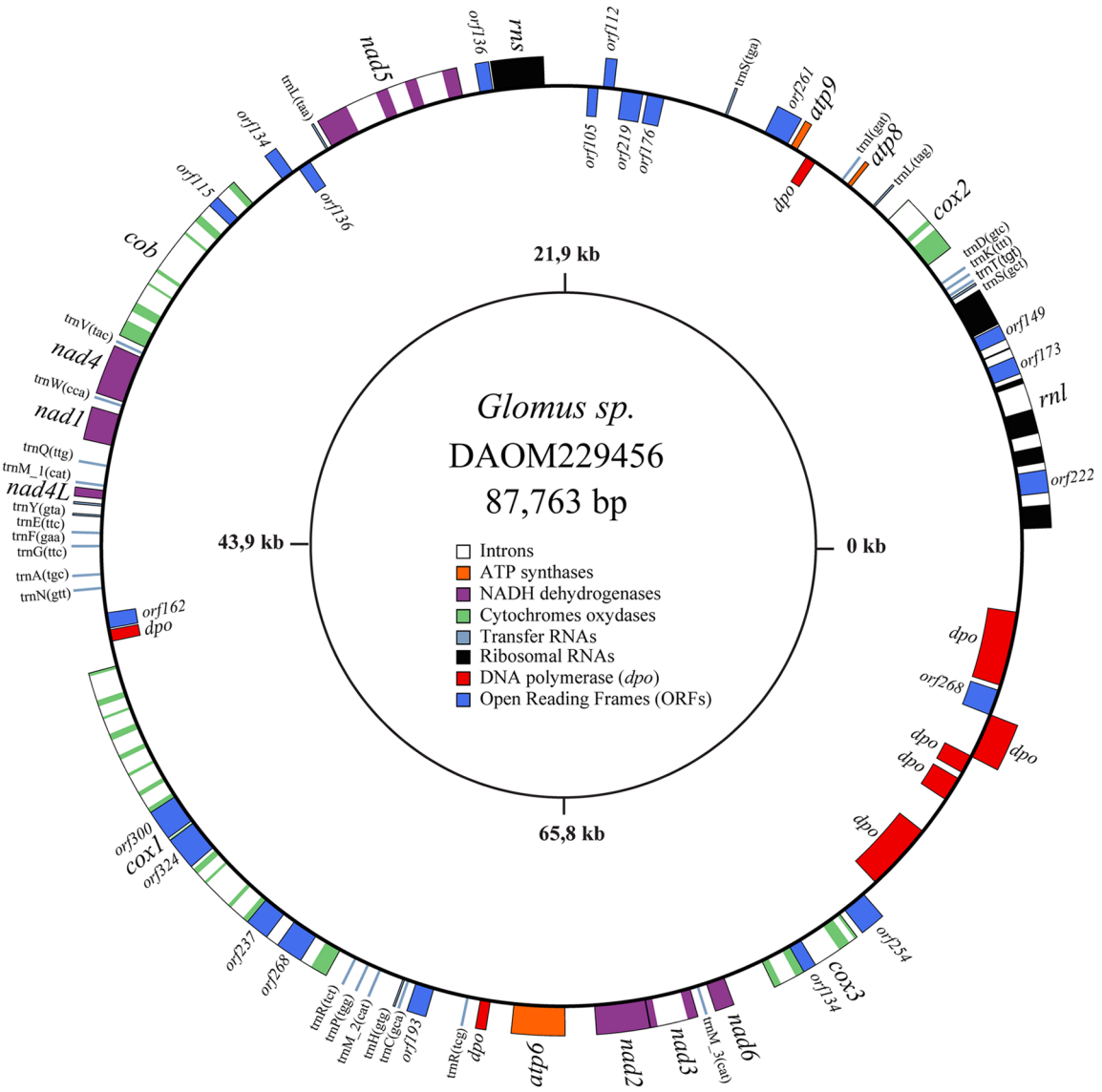


Figure 1. The *Glomus sp.* 229456 mitochondrial genome circular-map was opened upstream of *rnl*. Genes on the outer and inner circumference are transcribed in a clockwise and counterclockwise direction, respectively. Gene and corresponding product names are *atp6, 8, 9,* ATP synthase subunit 6; *cob*, apocytochrome b; *cox1–3*, cytochrome c oxidase subunits; *nad1–4, 4L, 5–6*, NADH dehydrogenase subunits; *rnl, rns*, large and small subunit rRNAs; A–W, tRNAs, the letter corresponding to the amino acid specified by the particular tRNA followed by their anticodon. Open reading frames smaller than 100 amino acids are not shown.
doi:10.1371/journal.pone.0060768.g001

disappear from mitochondrial genomes. Plasmid-related *dpo* insertions have been reported in the AMF *Gigaspora rosea*, but are virtually absent from the closely related paraphyletic zygomycetes. The mobility of mORFs, elements that thrive in *Glomus*, is mediated by the site-specific DNA endonuclease they encode. This endonuclease cleaves ORF-less alleles by creating a double-strand break in DNA and initiates the insertion and fusion of the mobile element. The same process, called *intron homing*, has been proposed for group I introns [43]. Several lines of phylogenetic evidence support the hypothesis of the evolutionary-independent ancestral origins of mORFs [44]. These highly mobile elements have the ability to carry group I introns [45], intergenic sequences [46], and coding sequences [47]. The first reported case of mitochondrial gene transfer caused by those elements was a mORF-mediated insertion of a foreign *atp6* carboxy-terminal in the blastocladiomycete *Allomyces macrogynus* [48].

The present study compared the mitochondrial genomes of the newly sequenced AMF species *Glomus sp.* DAOM226456 (a *Glomus diaphanum* like species based on spore morphology) with two isolates of the closely related *G. irregulare*. Along with a highly divergent intron insertion pattern, we found insertions of plasmid-related DNA polymerase and propagation of mobile open reading frame (mORFs) encoding endonucleases in *Glomus* mtDNAs. Our findings have brought to light the first evidence of AMF interspecific exchange of mitochondrial coding sequences entailing formation of gene hybrids in *Glomus sp. atp6*, *atp9* (coding for the subunit 6 and 9 of the ATP synthetase complex), *cox2* (cytochrome C oxidase subunit 2) and *nad3* (NADH dehydrogenase subunit 3) genes.

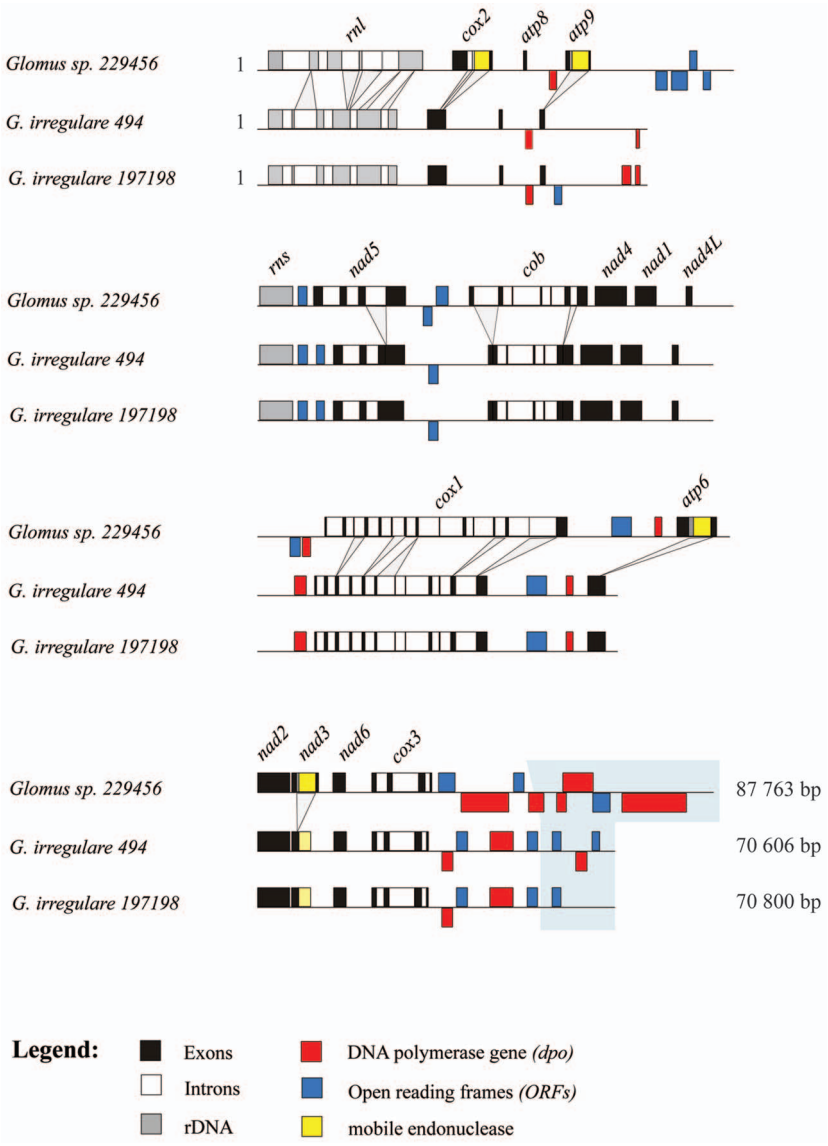


Figure 2. Comparative view of the three mitochondrial genomes linear map where the exons (black), introns (white), rDNA (gray), *dpo* plasmid insertions (red), ORFs (blue) and mobile endonuclease (yellow) are represented. Divergence in intron insertion pattern is indicated by projections. A hyper-variable region in the *cox3-rnl* intergene is boxed in grayscale.
doi:10.1371/journal.pone.0060768.g002

Table 1. Gene and intron content in AMF and selected fungal mtDNAs.

	Genes										
Species	*rnl, rns*	*atp 6, 8, 9*	*cob*	*cox 1, 2, 3*	*nad 1–6* [a]	*trn A–W*	*rnpB*	*rps3*	*ORFs* [b]	*Intron I* [c]	*Intron II* [c]
Glomus Sp. 229456	2	3	1	3	7	25	0	0	19	31	1
Glomus irregulare 494	2	3	1	3	7	25	0	0	8	26	0
Glomus irregulare 197198	2	3	1	3	7	25	0	0	8	26	0
Gigaspora rosea	2	3	1	3	7	25	0	0	4	13	1
Smittium culisetae	2	3	1	3	7	26	1	1	3	14	0
Mortierella verticillata	2	3	1	3	7	28	1	1	7	4	0
Rhizopus oryzae	2	3	1	3	7	23	1	0	4	9	0
Allomyces macrogynus	2	3	1	3	7	25	0	1	4	26	2
Saccharomyces cerevisiae [d]	2	3	1	3	0	25	1	1	3	9	4

[a]Includes *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5* and *nad6*.
[b]Only ORFs greater than 100 amino acids in length are listed, not including intronic ORFs and *dpo* and *rpo* fragments.
[c]Intron I and Intron II denote introns of group I and group II, respectively.
[d]*S. cerevisiae* strain FY 1679 [57].
doi:10.1371/journal.pone.0060768.t001

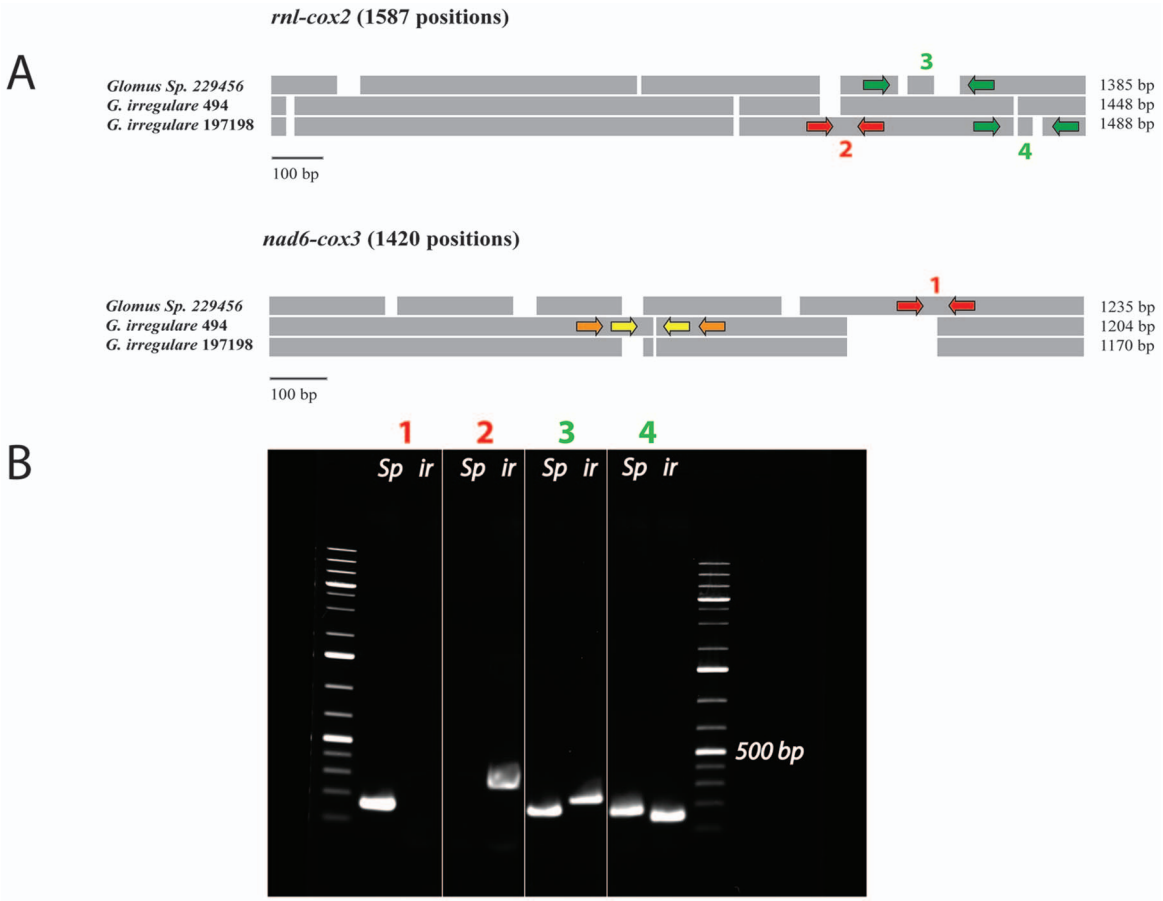


Figure 3. A) Schematic alignment representation of two mitochondrial intergenic regions (*rnl-cox2* and *cox3-nad6*) showing the presence of numerous insertions and deletions (indels). The red arrows indicate the approximate position of the PCR primers that yield strain-specific markers, while the green arrows indicate the position of PCR primers that produce a size-specific marker. The yellow and orange arrows indicate potential regions to design, respectively, specific and size-specific markers in *G. irregulare* 494. **B**) Agarose gel electrophoresis figure showing the PCR results of the proposed markers on *Glomus sp.* 229456 (*Gs*) and *G. irregulare* DAOM197198 (*Gi*) respectively for each marker. Marker 1 shows the *Glomus sp.* specific amplification (156 bp), while marker 2 shows the *G. irregulare* 197198 specific marker (263 bp). The size-specific marker 3 yield a length of 160 bp for *Glomus sp.* and 226 bp for *G. irregulare* 197198. Finally, the size-specific marker 4 yield a length of 159 bp for *Glomus sp.* and 131 bp for *G. irregulare* 197198.
doi:10.1371/journal.pone.0060768.g003

Materials and Methods

Fungal material

Spores and mycelium of *Glomus sp.* (DAOM-229456) and *G. irregulare* (DAOM 197198) were cultivated *in vitro* on a minimal (M) medium with carrot roots transformed with *Agrobacterium rhizogenes*, as described in the literature [49]. The medium was liquefied using a 0.82 mM sodium citrate and 0.18 mM citric acid extraction buffer solution. The resulting fungal material was further purified by hand under a binocular microscope, to remove root fragments.

DNA extraction

Spores and mycelium were suspended in 400 µL of the DNeasy Plant Mini Kit AP1 buffer (Qiagen) and crushed with a pestle in 1.5 ml microtubes, and the DNA was purified according to the manufacturer's recommendations. Purified DNA in a final elution volume of 40 µL was stored at −20°C until use.

RNA extraction

Fresh *Glomus sp.* fungal material was harvested from *in vitro* cultures. RNA extraction was performed using an E.Z.N.A. Fungal RNA Kit (Omega Biotek) according to manufacturer's recommendations. Total RNA was treated with Turbo DNase (Applied Biosystems) for 30 min at 37°C to remove residual DNA fragments that could interfere with downstream applications. In order to prevent chemical scission of the RNA during heat inactivation of the DNase at 75°C for 15 min, EDTA was added at a final concentration of 15 mM. In total, 40 µl of 100 ng/µl RNA was collected and stored at −80°C until use. The RNA concentration was determined using a Nanophotometer Pearl (Implen).

cDNA synthesis

From the total RNA previously extracted, 500 ng were used for cDNA synthesis with the SuperScript III reverse transcriptase kit (Life Technologies, Canada) according to manufacturer's recommendations, using oligo dT. The only change from these recommendations was the addition of $MgCl_2$ to a final concentration of 15 mM to compensate for the EDTA added in the previous step. In order to remove RNA complementary to the cDNA, 1 µl of RNase ONE ribonuclease (Promega, Canada) was added to the cDNA and incubated at 37°C for 20 min. The resulting cDNA was stored at −20°C until use.

Polymerase chain reaction (PCR)

The proposed intergenic markers to discriminate between *G. irregulare* DAOM197198 and *Glomus sp.* were tested by PCR using the KAPA2G Robust Hotstart ReadyMix PCR kit (KapaBiosystems, Canada). The specific primers used were respectively *rml-cox2*_197198_spec_F (5′-AAAGGAATTACATCGATTTA-3′), *rml-cox2*_197198_spec_R (5′- ACAAGAAGGTTTG-CATCGCTA-3′), *nad6-cox3*_dia_spec_F (5′- CCACTAGT-TAAGCTACCCTCTA-3′) and *nad6-cox3*_dia_spec_R (5′- AAT-CATACCGTGTGAAAGCAAG -3′). The variable length primers were *rml-cox2*_197198_size_F (5′- TAGGGATCAG-TACTTTAGCCAT -3′), *rml-cox2*_197198_size_R (5′-TCCTTACGGTATGAATGGTAAG -3′), *rml-cox2*_dia_size_F (5′- AGACTTCTTCAGTTCCACAATCA -3′) and *rml-cox2*_-dia_size_R (5′- ATGGCTAAAGTACTGATCCCTAC -3′). For 40 µl of PCR reaction volume, 12 µl of water, 20 µl of 2× PCR buffer, 3.5 µl of (5 µM) forward and reverse primers, and 1 µl of DNA were added. Cycling parameters were 94°C/3 min, followed by 38 cycles of: 94°C/30 sec, 54°C/25 sec, 72°C/45 sec and a final elongation at 72°C. PCR products were separated by electrophoresis in a 1.5% (w/v) agarose gel and visualized with GelRed under UV light.

Reverse transcriptase – polymerase chain reaction (RT-PCR)

Our objective with regard to PCR reactions on cDNA was to assess which regions of the gene hybrid reported in *atp6*, *atp9*, *cox2* and *nad3* were expressed at the mRNA level. For each of the four hybrids, a forward primer designed in the conserved 'core' structure of the gene (*atp6*_core_F: 5′-AGAGCAGTTTGA-GATTGTTAAG-3′, *atp9*_core_F: 5′-CTGGAGTAGGAG-TAGGGATAGT-3′, *cox2*_core_F: 5′-CATGGCAATTAG-GATTTCAAGA-3′ and *nad3*_core_F: 5′-TCGTTCCTTTGTTCGTGCTA-3′) was used in combination with three reverse primers designed respectively in the inserted C-terminal (*atp6*_insert_R1: 5′-AGCCTGAATAAGTGCAACAC-3′, *atp9*_insert_R1: 5′- GTAAGAAAGCCATCATGAGACA-3′, *cox2*_insert_R1: 5′-TGAGAAGAAAGCCATAACAAGT-3′ and *nad3*_insert_R1: 5′-AGAAGTATGAAAACCATAGCAATC-3′), the mobile ORF (*atp6*_mORF_R2: 5′-AGTCTTCGAATA-TACTGGCAG-3′, *atp9*_mORF_R2: 5′-TGTCGAGTCTC-CAAAGTATGT-3′, *cox2*_mORF_R2: 5′-ACT-GAATTCCTGTGTTTCGATCT-3′ and *nad3*_mORF_R2: 5′-TGACGAATGGTTAGACGATGT-3′) and the native C*-terminal portion of the corresponding gene (*atp6*_native_R3: 5′-CGTACCGTCGTAACAAGTAGA-3′, *atp9*_native_R3: 5′-CCATCATTAAGGCGAATAGA-3′, *cox2*_native_R3: 5′-CTAACAAACTCCCGACTATTACCT-3′ and *nad3*_native_R3: 5′-AGAATGAAGACCATTGCAAC-3′). To verify that there was no residual mitochondrial DNA in the cDNA, the primers Ctrl_positive_*nad5*exon4_689F (5′-ACCATTCTGTTATGTTC-TAATGT-3′) and Ctrl_positive_*nad5*exon4_689R (5′-GTCTGACTTAGCAGGTTAGTTAAG-3′) were designed in *nad5* exon 4 and used as a positive control on cDNA and negative control on RNA. The RT-PCR reactions were carried out using the KAPA2G Robust Hotstart ReadyMix PCR kit (KapaBiosystems, Canada) as described above in the PCR section.

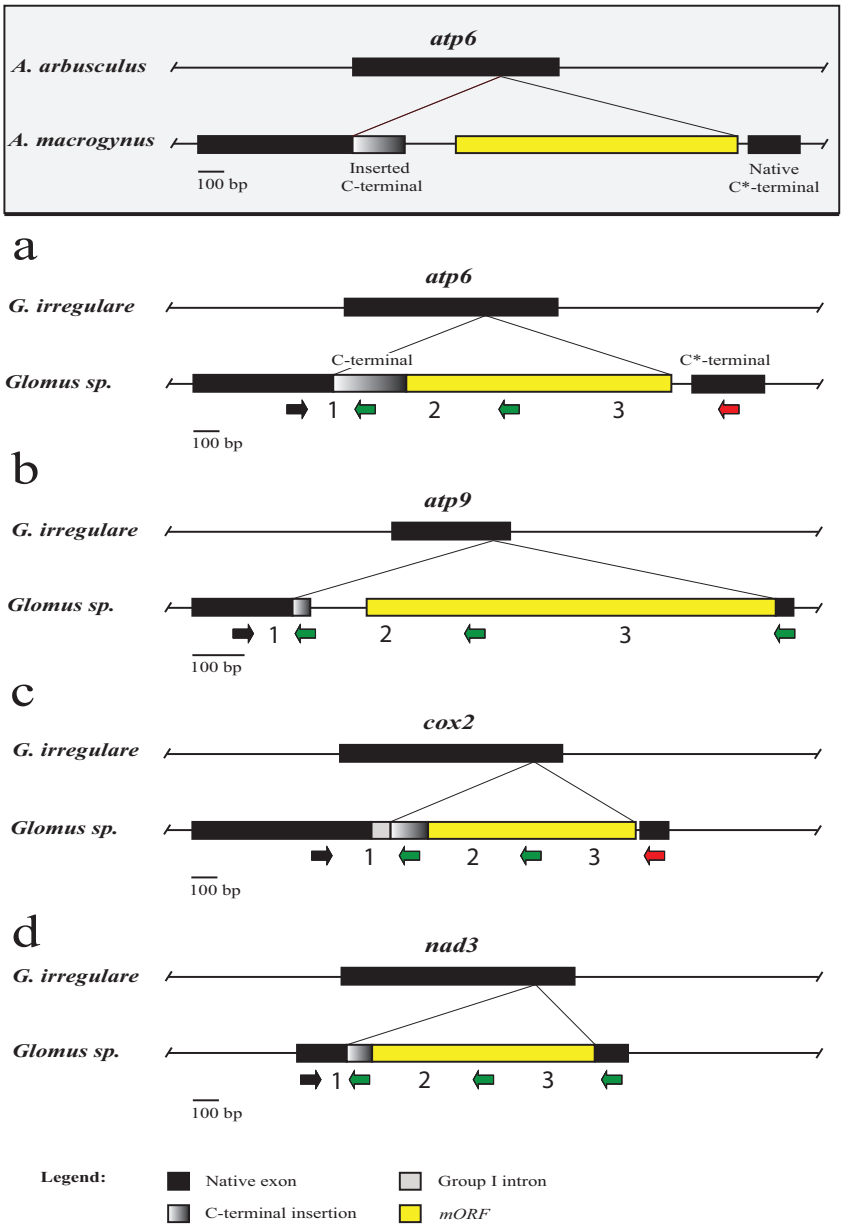
Cloning

Cloning reactions were performed on each successful RT-PCR amplification. The ligation reactions were done using the pGEM-T Easy Vector Systems kit (Promega, Canada) according to manufacturer's recommendations. The transformation was carried out in *E. coli* DH5 alpha competent cells. Bacterial colonies were screened via PCR using T7 and SP6 universal primers as described in the PCR section.

Sequencing, assembly and gene annotation

Glomus sp. total DNA was sequenced using 454 Titanium Flex shotgun technology (one plate) and the respective resulting 1,078,190 reads were assembled with Newbler (Genome Quebec Innovation Center, McGill University, Montreal, Canada). Gene annotation was performed with MFannot (http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl), followed by manual inspection and introduction of missing gene features as described in Nadimi et al., (2012). *G. irregulare* isolates 494 and DAOM-197198 mtDNAs (accession numbers FJ648425 and HQ189519 respectively) were used for comparison. Sequencing of the cloned RT-PCR products was performed on the same sequencing platform, using Sanger technology with T7 and SP6 universal primers.

A


atp6
A. arbusculus
A. macrogynus
100 bp
Inserted
C-terminal
Native
C*-terminal
a
atp6
G. irregulare
Glomus sp.
C-terminal
C*-terminal
1
2
3
100 bp
b
atp9
G. irregulare
Glomus sp.
1
2
3
100 bp
c
cox2
G. irregulare
Glomus sp.
1
2
3
100 bp
d
nad3
G. irregulare
Glomus sp.
1
2
3
100 bp
Legend:
Native exon
Group I intron
C-terminal insertion
mORF

B

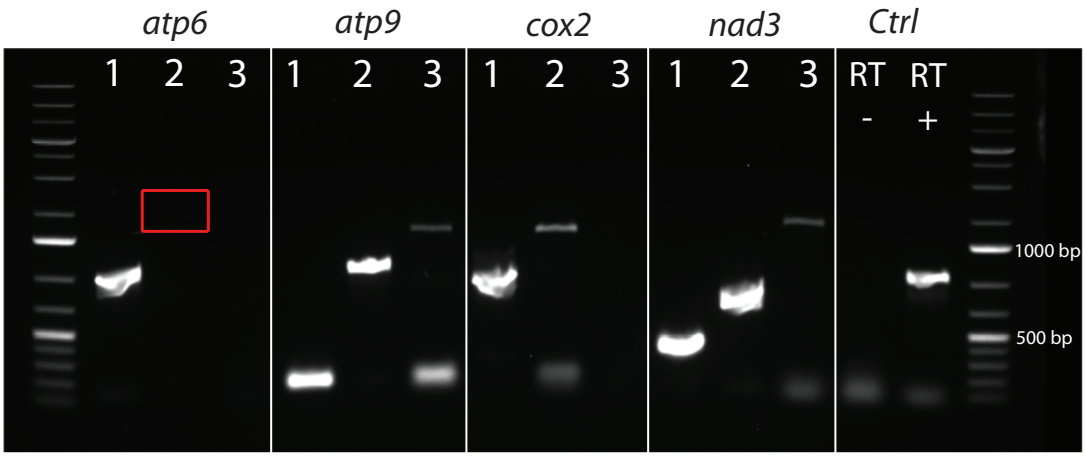

atp6
atp9
cox2
nad3
Ctrl
1 2 3
1 2 3
1 2 3
1 2 3
RT RT
- +
1000 bp
500 bp

Figure 4. Comparison of gene hybrids atp6, atp9, cox2 and nad3. A) The *atp6* gene hybrid reported for *Allomyces macrogynus* (grayscale, boxed) is used as a reference in a comparison of the most similar *atp6* (a), *atp9* (b), *cox2* (c) and *nad3* (d) genes in *Glomus sp.* mtDNA. Each occurrence is put in perspective with the gene of a close relative (either *Allomyces arbusculus* or *G. irregulare*) in order to show the insertion point of the foreign element with the projections. Exons are in black, while the inserted foreign C-terminal is shaded in gray. The mobile endonuclease element is in yellow. For each gene, the black arrow indicates the position of the forward primer used in the downstream RT-PCR experiment in combination with three different reverse primers. The green arrows indicate expression at the RNA level of the corresponding portion of the gene, while the red arrows indicate a negative amplification. **B**) Agarose gel electrophoresis figure showing the RT-PCR results. For each gene hybrid, the expression at the RNA level was tested using a forward primer in the conserved gene core and a reverse primer respectively in the inserted C-terminal (1), the mobile endonuclease (2) and the native C*-terminal portion (3). Primers in *nad5* exon 4 were used as a positive control on cDNA (RT +) and negative control on RNA (RT −). The expected size of the amplified fragments was: *atp6* inserted C-terminal (684 bp), *atp9* inserted C-terminal (149 bp), *atp9* mORF (717 bp), *atp9* native C*-terminal (1085 bp), *cox2* inserted C-terminal (938 bp), *cox2* mORF (1291 bp), *nad3* inserted C-terminal (261 bp), *nad3* mORF (597 bp), *nad3* native C*-terminal (1183 bp) and the positive control in *nad5* exon 4 had an expected amplicon size of 689 bp. The red box indicates a faint band that is present on the gel.
doi:10.1371/journal.pone.0060768.g004

Phylogenetic analysis

For each gene of interest (*atp6*, *atp9*, *cox2* and *nad3*) in 12 AMF species, the dataset contains the corresponding C*-terminal for: *Glomus sp.* DAOM-229456, *G. irregulare* isolate 494, *G. irregulare* DAOM-197198, *G. irregulare* DAOM-240415, *G. irregulare* DAOM-234179, *G. irregulare* DAOM-234328, *G. irregulare* DAOM-213198, *Glomus sp.* DAOM-240422, *G. fasciculatum* DAOM-240159, *G. aggregatum* DAOM-240163, *G. cerebriforme* DAOM-227022, *Gigaspora rosea* DAOM-194757 (accession number JQ693396) and 3 selected fungal representatives: *Mortierella verticillata* (accession number AY863211), *Smittium culisetae* (accession number AY863213) and *Rhizopus oryzae* (accession number AY863212). The sequences were deposited in databases under the accession numbers: JX074786-JX074817. The reference phylogeny was constructed using the concatenated 'core' sequence (without the C*-terminal portion used previously) of the same four genes. The DNA sequence alignments and the inference of maximum likelihood trees using GTR+G (with five distinct gamma categories) were performed using the integrated program MEGA version 5 [50]. Bootstrap resampling (1000 replicates) was carried out to quantify the relative support for each branch of the trees. Bayesian analysis were done using MrBayes version 3.2 using the GTR+G model (with five distinct gamma categories), four independant chains, one million cycles, tree sampling every 100 generations and a burn-in value of 25%.

Results and Discussion

Glomus sp. genome organization and structure

The complete sequence of the *Glomus sp.* 229456 mt genome was a double-stranded circular DNA molecule, exempt of polymorphism, with a size of 87,763 bp. The annotated sequence of *Glomus sp.* was deposited in GenBank under the accession number JX065416. Its mtDNA harbors the typical set of 41 mitochondrial genes found in other AMF (two rRNAs, 14 protein coding genes (PCGs) and 25 tRNAs). The PCGs include three ATP synthetase (*atp*), one cytochrome b (*cob*), three cytochrome C oxydase (*cox*) and seven NADH dehydrogenase (*nad*) genes. Also, 19 ORFs and 31 introns are inserted in this newly sequenced mt genome (Figure 1).

Comparative view of three *Glomus* mtDNAs

The gene content in *Glomus sp.* and *G. irregulare* mitochondrial genomes is similar to that found in zygomycetes, except for *rps3* and *rnpB*. The mtDNAs of both AMF species have the same gene order, and all genes are transcribed from one strand with very similar coding regions except for the insertion of mobile ORF elements (mORFs) in the *atp6*, *atp9*, *cox2* and *nad3* genes of *Glomus sp.* (Figure 2). However, there are many differences in the number of introns, some of which carry more substantial sequence differences than do the coding sequences. The differences in the presence of introns and mORFs explain the inflated genome size of 87,763 bp in *Glomus sp.*, as compared to 70,800 bp in *G. irregulare* 197198 (Table 1). *Glomus sp. cox1* intron 8 is the homolog of an intron inserted at the same position in *Rhizopus oryzae* and angiosperms (with 76 and 79% of sequence identity, respectively) [28]. *G. irregulare cox1* intron 7 is also inserted at the same position, but has an eroded ORF encoding the homing endonuclease gene, and thus also shares identity with the intron RNA secondary structure of *R. oryzae* and plants. The plant *cox1* intron was thought to have been acquired from a fungal donor, due to the proximity of its clade to that of fungi rather than to the non-vascular plant *Marchantia*. Knowing the extent to which the intron has spread in angiosperms [51,52], it would be interesting to see whether such an invasion has also occurred within the Glomeromycota phylum.

Table 2. Description of the gene hybrids found in *Glomus sp.* 229456 mtDNA.

	atp6	*atp9*	*cox2*	*nad3*
Total length	1569	1171	1894	1242
CDS length	1569	225	837	1242
Features				
Group I intron	-	-	[684–922]	
Inserted C-terminal	[537–774] [1]	[175–225] [1]	[923–1018] [1]	[194–313] [1]
mORF	[550–1567] [1]	[334–1119] [1]	[1187–1738] [1]	[451–867] [1]
Native C*-terminal	[1582–1860]	[1120–1171] [1]	[1739–1894]	[1116–1238] [1]
Remarks	C-terminal and m*ORF* in phase with native gene	C-terminal in phase with native gene.	Partial inserted C-terminal in phase with native gene.	C-terminal and m*ORF* in phase with native gene.

[1]Gene hybrid features that are expressed at the mRNA level (see Figure 4B).
doi:10.1371/journal.pone.0060768.t002

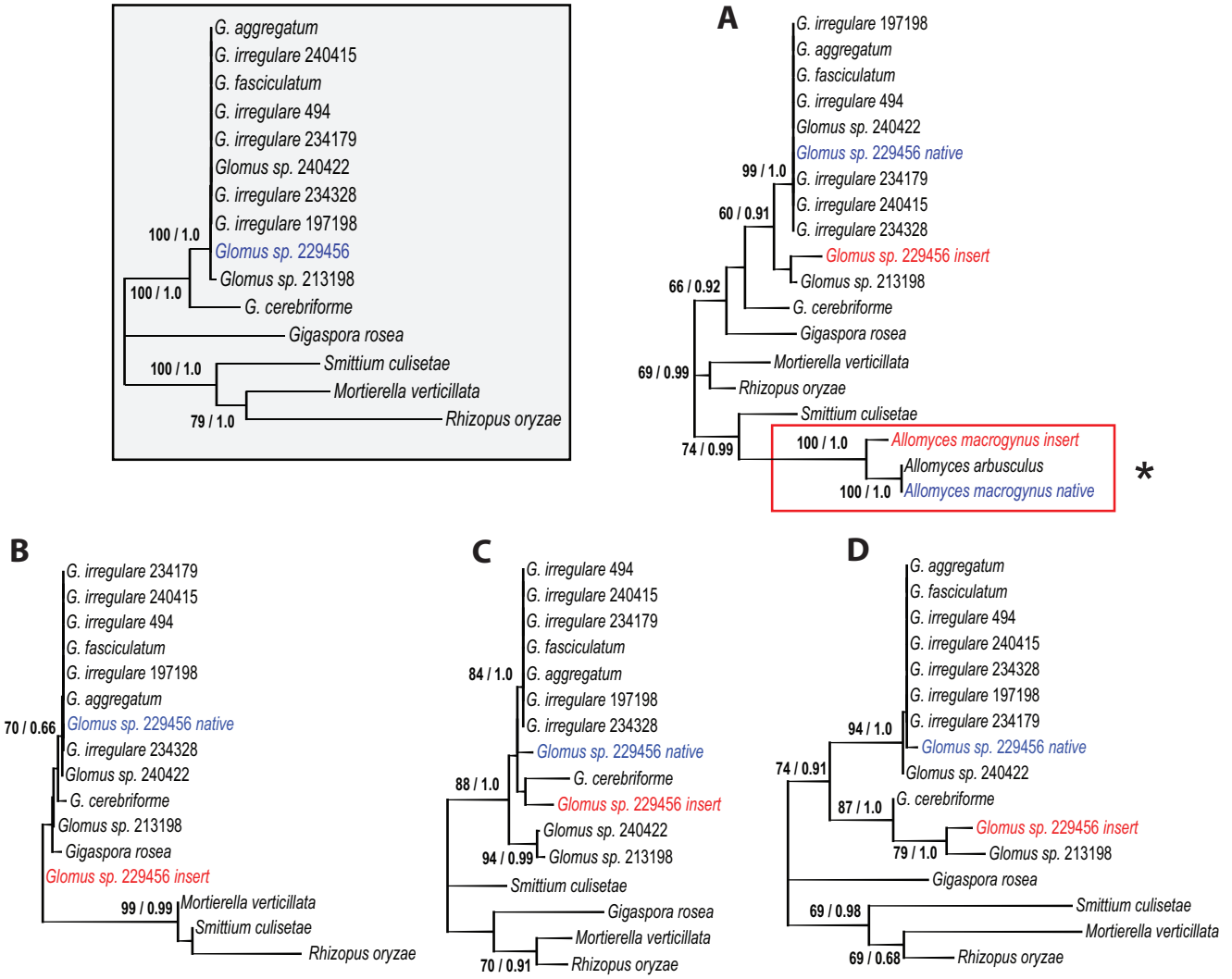


Figure 5. Unrooted maximum likelihood trees obtained with the GTR+G model (5 distinct gamma categories). The first number at branches indicates ML bootstrap values with 1000 bootstrap replicates and the second number indicates posterior probability values of a MrBayes analysis with four independant chains. Bayesian inference predict similar trees (not shown). The concatenated tree of the *atp6*, *atp9*, *cox2* and *nad3* 'core' genes (without the duplicated C*-terminal portion) (1489 alignment positions) of selected AMF representatives (grayscale boxed) are compared with those of the *atp6* (298 alignment positions), where the red box with the asterisk point out to the reference *Allomyces spp.* HGT event (Figure 4) (A), *atp9* (51 alignment positions) (B), *cox2* (106 alignment positions) (C) and *nad3* (120 alignment positions) (D) C*-terminals. The *Glomus sp.* native C*-terminals are in blue, while the inserted C-terminals are in red.
doi:10.1371/journal.pone.0060768.g005

Further, intergenic regions differ substantially in sequence: some are identical while others show signs of very fast, substantial changes including point mutations, insertions, deletions and inversions (Figures 2 and 3). Most of these differences occur in the *cox3-rnl* intergene, a large hyper-variable region that has been invaded by *dpo* fragments. The variations observed in intergenic regions provide an opportunity to develop species- specific molecular markers as shown in Figure 3, and even isolate-specific markers or methods allowing reliable identification and/or quantification of these fungi. Lack of efficient and powerful molecular markers for AMF identification and quantification constitutes a major problem that limits the analysis of population genetics and field studies in AMF. Mitochondrial DNA is homogeneous within the AMF individuals studied to date, but evidence of genetic polymorphism between *G. irregulare* isolates has been observed in intergenic regions. They harbor highly conserved genes as well as highly variable regions, which promises to facilitate AMF barcoding at different taxonomic levels, an analysis that is currently challenging to carry out using nuclear genes. Hyper-variable intergenic regions with eroded *dpo* insertions and indels in intergenic regions constitute useful mitochondrial areas on which to focus attention in order to develop suitable markers for discriminating isolates of the same species. Intron insertion pattern variations, genome reorganizations (such as gene shuffling) and coding region divergences will make it possible to distinguish between different AMF species, genera and families.

Our 454 pyrosequencing data and direct PCR sequencing showed that *G. irregulare* DAOM197198 and *Glomus sp.* mtDNAs are homogeneous, meaning that all the mitochondrial genomes in a given isolate are essentially identical, in stark contrast to the nuclear genomes. Our results confirm the previous report by Lee *et al.* (2009) suggesting homoplasmy in the first completed Glomeromycota mitochondrial genome of the AMF *G. intraradices* (*G. irregulare* isolate 494). A rapid and effective mitochondrial

segregation mechanism was suggested to explain those findings. It was previously demonstrated that isolates of the same species can exchange nuclear material through anastomosis [10], but exchange of divergent mitochondrial haplotypes has yet to be shown. This leads us to question whether polymorphism does indeed occur through anastomosis, and for how many generations mitochondrial heteroplasmy is maintained.

Rapid expansion of plasmid-like DNA polymerase sequences in *Glomus*

Plasmid-related DNA polymerase genes are found in mobile mitochondrial plasmids that occur either as free linear or circular DNAs, and have been shown to also insert into mtDNA (for review see [35]). One striking feature in the comparison of the two closely related *Glomus* species is the presence of numerous *dpo* insertions in the intergenic regions of their mtDNA (Figure 2). All three *Glomus* mtDNAs contain a large number of *dpo* fragments, most of which are substantially divergent in sequence and therefore are most likely the result of independent plasmid insertion events. Even the two *G. irregulare* (isolates 494 and DAOM197198), otherwise almost identical in sequence, differ in *dpo* sequence, which supports the interpretation that *dpo* insertion occurs repeatedly and frequently through evolutionary time. A *bona fide* and complete *dpo* gene is present in *Glomus sp.*, and its sequence is different from those in *G. irregulare* isolates. Because of its complete length, it most likely results from a recent insertion event. There is no evidence that *dpo* is functional when inserted in mtDNA. As in numerous other cases, *dpo* coding regions are fragmented in *Glomus* and occur on both strands, representing a good indicator of a genomic region experiencing little if any selective evolutionary constraints.

The source of the *dpo* insertions in *Glomus* mtDNA remains elusive. We did not find any free mitochondrial plasmids in our *Glomus sp.* and *G. irregulare* isolate DAOM197198 shotgun data (combining nuclear and mitochondrial DNAs), as we did for *Gigaspora rosea*, where a 3582 bp contig with high sequence coverage was found [34]. However, since the *Glomus* strains used in this study come from aseptic *in vitro* cultures, and even though the *G. rosea* fungal material was extracted from *in vivo* greenhouse pot cultures, we cannot rule out the possibility that an environmental vector is the source of *dpo* plasmids and is responsible for their propagation in *G. rosea*. Interestingly, *dpo* plasmids have been found to occur in numerous plants, notably in *Daucus carota* [53] which is used as a host plant for AMF cultures *in vitro*. The obligate biotrophic dependence of AMF on plants could be one of the reasons that *dpo* insertions are most abundant in *Glomus* mtDNAs yet virtually absent in mitochondrial sequences of the Blastocladiomycota (except for a single 100 amino acid long fragment occurrence in *Smittium culisetae* mtDNA), which is the closest phylogenetic group to the Glomeromycota.

Mobile element insertions have been shown to trigger genomic rearrangements such as gene shuffling through homologous recombination [54] and even genome linearization [35,55,56]. Whenever sequence repeats occur, more than one genome conformation may exist, but we have no evidence that this happens in *Glomus* mtDNA. It would be interesting to examine whether numerous recent *dpo* insertions with high sequence similarity might act as genomic repetitions and give rise to genome reorganization in closely related AMF species. Integrated plasmid segments within mitochondrial genomes, even though they are neutral or cryptic, could promote genomic rearrangements.

Mobile ORF elements (mORFs) in *Glomus*

Although most ORF-encoding endonuclease genes are inserted in introns where they have been shown to play a role in propagation, they can also be present in genes in which their evolutionary impact is less obvious. We identified numerous mORFs encoding endonuclease genes unique to *Glomus sp.* isolate DAOM-229456 mtDNA. When we annotated the sequences of the *atp6*, *atp9*, *cox2* and *nad3* genes, we observed that they all have a peculiar organization. Indeed, these genes harbor a carboxy-terminal duplication (C*-terminal) that was found downstream of a mORF insertion. For example, in the *atp6* gene, the duplicated portion of the C-terminal was found about 1000 bp downstream, following an inserted LAGLIDADG endonuclease ORF. When we compared the DNA sequence of the C*-terminal portion with the corresponding sequence of *G. irregulare* isolate 494, a close relative to *Glomus sp.*, we found a 91.2% nucleotide identity. In contrast, the comparison between the *Glomus sp. atp6* duplicated carboxy-terminals (C-terminal and C*-terminal) showed a low sequence identity of 63.5%. Interestingly, comparison of the *Glomus sp.* C*-terminal amino acid sequences with the corresponding portion in *G. irregulare* showed 100% identity, indicating that the mutations observed in DNA are all synonymous. However, the comparison of the amino acid sequences of *Glomus sp. atp6* carboxy-terminals showed 91% identity.

Surprisingly, when we designed a forward primer in the upstream sequence (5′ gene portion) and two reverse primers in the C-terminal and C*-terminal respectively, we found that the C-terminal is transcribed with the upstream sequence resulting in a putative hybrid transcript while the C*-terminal was not expressed into mRNA. Thus we hypothesized that the C-terminal portion could have been acquired from a donor through horizontal gene transfer (HGT). We also observed similar organization in *atp9*, *cox2* and *nad3* genes of *Glomus sp.* where the carboxy-terminal portion (C*-terminal) was replaced partially or completely by one carried by a mORF (C-terminal) encoding a LAGLIDADG endonuclease (except in *atp9*, a GIY-YIG family endonuclease) (Figure 4A: a, b, c and d). In *atp6*, the insert lacks a stop codon and the ORF is in phase with the native gene. In *atp9*, the insert, along with the mORF, completely replaces the native 3′ end, while in *cox2* and *nad3* only a portion of the carboxy-terminal is replaced (Table 2). The resulting gene hybrids are expressed at the mRNA level in all four cases as shown in Figure 4B. After sequencing of the cDNA bands, we found that the mORF and the inserted C-terminal are integral parts of the transcript in all four genes. However, in *atp6* and *cox2*, the native C*-terminal was not expressed into mRNA.

These gene hybrid structures are similar to that of the *atp6* gene previously described in the *Allomyces macrogynus* (Figure 4, grayscale box), a species that belongs to the basal fungal phylum Blastocladiomycota [48]. The same scenario has also been observed in the *Rhizopus oryzae atp9* and *Mortierella verticillata cox2* genes [30]. These hybrids contain a carboxy-terminal duplication as well as a mORF encoding an endonuclease, which has been biochemically demonstrated to be responsible for the element mobility. In *Allomyces macrogynus*, the inserted C-terminal was shown to have been recently acquired by HGT based on the divergence in sequence it had with the native C*-terminal, while the latter had a perfect sequence identity with the corresponding gene portion of the closely related species *Allomyces arbusculus*.

The *Glomus sp. atp6*, *atp9*, *cox2* and *nad3* native C*-terminals showed higher nucleotide sequence identity to those of *G. irregulare* 494 (91, 98, 93 and 98%, respectively) than their duplicated C-terminal counterparts (64, 71 and 81 and 73%, respectively) (Figures S1, S2, S3, and S4 and Tables S1, S2, S3, and S4). However at the protein level, the comparison of the C*-terminal

amino acid sequences of the *atp6*, *atp9*, *cox2* and *nad3* genes with the corresponding portion in *G. irregulare* 494 was 100% for *atp6*, *94% for atp9, and* 100% for *cox2* and *nad3*. The high sequence identity of the native *Glomus sp.* C*-terminals with *G. irregulare 494*, is in stark contrast to the low similarity observed with the inserted C-terminal portions and points to a recent HGT event, as was described in *Allomyces spp.* [*48*]. However, the HGT hypothesis could likely apply to the *atp6* and *cox2* genes, since their native C*-terminal portion is no longer translated and could undergo rapid divergence. For the *atp9* and *nad3* hybrids, even though it is less parsimonious, the observed sequence divergence between the duplicated portions could have been caused by independent evolution following the mobile element insertion, since both are expressed in the mRNA transcript. It would also be interesting to see if some of the reported gene hybrids can still accomplish their functions at the protein level, given that the mORF and both C*-terminals are expressed in some cases. They are apparently expressed pseudogenes but post-translational modification mechanisms may be in place to ensure that the resulting protein is functional. We did not find a mORF-less copy of those genes that could have been transferred to the *Glomus sp.* nuclear genome that could explain a pseudogenization in *Glomus sp.* mtDNA.

In regards to the HGT hypothesis, and in order to evaluate whether there is a plausible donor for the duplication, we compared the carboxy-terminal sequence of these genes with those in 11 *Glomus spp.* (to avoid redundancy we didn't add the *G. margarita* sequences since they are identical to *G. rosea*) and three phylogenetically related fungal representatives (Figure 5). In all four *Glomus sp.* gene hybrids (*atp6*, *atp9*, *cox2* and *nad3*), the native C*-terminal sequences cluster within the *Glomus spp.* group as expected given the reference phylogeny (Figure 5, grayscale box), thereby supporting a recent insertion of the foreign element. The *atp6* gene carboxy-terminal comparison (Figure 5A) shows that the mORF-derived C-terminal is related to a *Glomus sp.* isolate DAOM213198 with a moderate 60% bootstrap value. Surprisingly, in *atp9* (Figure 5B) the inserted C-terminal is even more distantly related to *Glomus spp.* than to *G. rosea*. In *cox2* (Figure 5C) the *Glomus sp.* inserted C-terminal and the more divergent AMF species *G. cerebriforme* are in the same cluster. Finally, the *nad3* C-terminal clustered with *Glomus sp.* 213198, as it was the case for *atp6*, with a 79% bootstrap value (Figure 5D). Also, the *nad3* gene shows high variability in length in *Glomus spp.*, due to the insertion of those elements.

In all four cases, the native *Glomus sp.* C*-terminal is nested within the *Glomus spp.* group and the inserted C-terminal is in a different cluster. Although it is difficult to pinpoint the donor of the sequence duplications, due to the possibly complex evolutionary history of those mobile elements with numerous insertion/loss events and 3′ end reshufflings, our data suggest HGT from a foreign AMF species, and thus the first reported occurrence in Glomeromycota. The presence of foreign DNA elements could potentially hamper mitochondrial gene phylogeny analysis unless the foreign C-terminals are carefully removed from the native portion of the gene.

Conclusion

The inclusion of mitochondrial sequences from phylogenetically distant AMF species in the database is essential for developing a better understanding and classification of AMF within fungi. The mitochondrial genome comparison presented here for two closely related AMF species reveals substantial changes in mitochondrial gene sequences, resulting from *dpo* plasmid insertions and mobile ORFs invasions, along with intergenic sequence variation. This illustrates the importance of adding closely related species to the numerous isolates of the same species in the AMF mitochondrial genome collection. Comparative mitochondrial genomics, together with a broader sequencing effort in AMF, opens new avenues for the development of molecular markers at different evolutionary distances. It would be interesting to identify the source of plasmid-related DNA polymerase in AMF mtDNA, which should provide an estimate of the extent to which it is present within the Glomeromycota phylum and an assessment of the consequences on mitochondrial genome organization. Also, the mORF-carried foreign C-terminal described here represents the first reported evidence of HGT in AMF. The intimate relationship between AMF, the roots of their plant symbiont and soil microorganisms might be a perfect biological context to facilitate such transfers. To what extent the mobilome and HGT may have contributed to AMF evolution is a topic that merits exploration in future studies.

Supporting Information

Figure S1 Multiple DNA sequence alignment of numerous AMF representatives of the atp6 native C-terminals along with the Glomus sp. 229456 putative foreign inserted C*-terminal.
(TIF)

Figure S2 Multiple DNA sequence alignment of numerous AMF representatives of the atp9 native C-terminals along with the Glomus sp. 229456 putative foreign inserted C*-terminal.
(TIF)

Figure S3 Multiple DNA sequence alignment of numerous AMF representatives of the cox2 native C-terminals along with the Glomus sp. 229456 putative foreign inserted C*-terminal.
(TIF)

Figure S4 Multiple DNA sequence alignment of numerous AMF representatives of the nad3 native C-terminals along with the Glomus sp. 229456 putative foreign inserted C*-terminal.
(TIF)

Table S1 Sequence identity matrix of the atp6 native C-terminals along with the Glomus sp. 229456 putative foreign inserted C*-terminal.
(DOC)

Table S2 Sequence identity matrix of the atp9 native C-terminals along with the Glomus sp. 229456 putative foreign inserted C*-terminal.
(DOC)

Table S3 Sequence identity matrix of the cox2 native C-terminals along with the Glomus sp. 229456 putative foreign inserted C*-terminal.
(DOC)

Table S4 Sequence identity matrix of the nad3 native C-terminals along with the Glomus sp. 229456 putative foreign inserted C*-terminal.
(DOC)

Acknowledgments

We thank Dr. B.F. Lang for bioinformatics assistance and access to an automated organelle genome annotation software and to Dr. David Morse and Dr. Terrence Bell for comments on the manuscript. We also like to

thank Biopterre centre du développement des bioproduits and CRBM for their help.

Author Contributions

Conceived and designed the experiments: MH. Performed the experiments: DB BI. Analyzed the data: DB MN. Contributed reagents/materials/analysis tools: DB MH. Wrote the paper: DB MH.

References

1. Wang B, Qiu YL (2006) Phylogenetic distribution and evolution of mycorrhizas in land plants. Mycorrhiza 16: 299–363.
2. Smith S, Read D (2008) Mycorrhizal Symbiosis. Cambridge: UK: Academic Press.
3. Azcón-Aguilar C, Barea JM (1997) Arbuscular mycorrhizas and biological control of soil-borne plant pathogens – an overview of the mechanisms involved. Mycorrhiza 6: 457–464.
4. St-Arnaud M, Vujanovik V (2007) Effects of Arbuscular Mycorrhizal Fungi on Plant Diseases and Pests Mycorrhizae in Crop Production: Applying knowledge. Binghamton, NY: Haworth Press.
5. Datnoff LE, Nemec S, Pernezny K (1995) Biological Control of Fusarium Crown and Root Rot of Tomato in Florida Using Trichoderma harzianum and Glomus intraradices. Biological Control 5: 427–431.
6. Cordier C, Pozo MJ, Barea JM, Gianinazzi S, Gianinazzi-Pearson V (1998) Cell Defense Responses Associated with Localized and Systemic Resistance to Phytophthora parasitica Induced in Tomato by an Arbuscular Mycorrhizal Fungus. Molecular Plant-Microbe Interactions 11: 1017–1028.
7. Pozo MJ, Cordier C, Dumas-Gaudot E, Gianinazzi S, Barea JM, et al. (2002) Localized versus systemic effect of arbuscular mycorrhizal fungi on defence responses to Phytophthora infection in tomato plants. Journal of Experimental Botany 53: 525–534.
8. Ismail Y, Hijri M (2012) Arbuscular mycorrhisation with Glomus irregulare induces expression of potato PR homologues genes in response to infection by Fusarium sambucinum. Functional Plant Biology 39: 236–245.
9. Ismail Y, McCormick S, Hijri M (2011) A Fungal Symbiont of Plant-Roots Modulates Mycotoxin Gene Expression in the Pathogen Fusarium sambucinum. PLoS ONE 6: e17990.
10. Croll D, Giovannetti M, Koch AM, Sbrana C, Ehinger M, et al. (2009) Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus Glomus intraradices. New Phytologist 181: 924–937.
11. Bever JD, Wang M (2005) Arbuscular mycorrhizal fungi: Hyphal fusion and multigenomic structure. Nature 433: E3–E4.
12. Angelard C, Sanders IR (2011) Effect of segregation and genetic exchange on arbuscular mycorrhizal fungi in colonization of roots. New Phytologist 189: 652–657.
13. Colard A, Angelard C, Sanders IR (2011) Genetic Exchange in an Arbuscular Mycorrhizal Fungus Results in Increased Rice Growth and Altered Mycorrhiza-Specific Gene Transcription. Applied and Environmental Microbiology 77: 6510–6515.
14. Corradi N, Croll D, Colard A, Kuhn G, Ehinger M, et al. (2007) Gene Copy Number Polymorphisms in an Arbuscular Mycorrhizal Fungal Population. Applied and Environmental Microbiology 73: 366–369.
15. Halary S, Malik S-B, Lildhar L, Slamovits CH, Hijri M, et al. (2011) Conserved Meiotic Machinery in Glomus spp., a Putatively Ancient Asexual Fungal Lineage. Genome Biology and Evolution 3: 950–958.
16. Tisserant E, Kohler A, Dozolme-Seddas P, Balestrini R, Benabdellah K, et al. (2012) The transcriptome of the arbuscular mycorrhizal fungus Glomus intraradices (DAOM 197198) reveals functional tradeoffs in an obligate symbiont. New Phytologist 193: 755–769.
17. Sanders Ian R (2011) Fungal Sex: Meiosis Machinery in Ancient Symbiotic Fungi. Current biology: CB 21: R896–R897.
18. Kuhn G, Hijri M, Sanders IR (2001) Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi. Nature 414: 745–748.
19. Hijri M, Sanders IR (2005) Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei. Nature 433: 160–163.
20. Pawlowska TE, Taylor JW (2004) Organization of genetic variation in individuals of arbuscular mycorrhizal fungi. Nature 427: 733–737.
21. Pawlowska TE, Taylor JW (2005) Arbuscular mycorrhizal fungi: Hyphal fusion and multigenomic structure (reply). Nature 433: E4–E4.
22. Boon E, Zimmerman E, Lang BF, Hijri M (2010) Intra-isolate genome variation in arbuscular mycorrhizal fungi persists in the transcriptome. Journal of Evolutionary Biology 23: 1519–1527.
23. VanKuren NW, den Bakker HC, Morton JB, Pawlowska TE (2013) Ribosomal RNA gene diversity, effective population size, and evolutionary longevity in asexual Glomeromycota. Evolution 67: 207–224.
24. Raab PA, Brennwald A, Redecker D (2005) Mitochondrial large ribosomal subunit sequences are homogeneous within isolates of Glomus (arbuscular mycorrhizal fungi, Glomeromycota) Mycological Research 109: 1315–1322.
25. Lee J, Young JP (2009) The mitochondrial genome sequence of the arbuscular mycorrhizal fungus Glomus intraradices isolate 494 and implications for the phylogenetic placement of Glomus. New Phytol 183: 200–211.
26. Börstler B, Raab PA, Thiéry O, Morton JB, Redecker D (2008) Genetic diversity of the arbuscular mycorrhizal fungus Glomus intraradices as determined by mitochondrial large subunit rRNA gene sequences is considerably higher than previously expected. New Phytologist 180: 452–465.
27. Thiéry O, Börstler B, Ineichen K, Redecker D (2010) Evolutionary dynamics of introns and homing endonuclease ORFs in a region of the large subunit of the mitochondrial rRNA in Glomus species (arbuscular mycorrhizal fungi, Glomeromycota). Molecular Phylogenetics and Evolution 55: 599–610.
28. Lang B, Franz, Hijri M (2009) The complete Glomus intraradices mitochondrial genome sequence – a milestone in mycorrhizal research. New Phytologist 183: 3–6.
29. Formey D, Moles M, Haouy A, Savelli B, Bouchez O, et al. (2012) Comparative analysis of mitochondrial genomes of Rhizophagus irregularis – syn. Glomus irregulare – reveals a polymorphism induced by variability generating elements. New Phytol 196: 1217–1227.
30. Seif E, Leigh J, Liu Y, Roewer I, Forget L, et al. (2005) Comparative mitochondrial genomics in zygomycetes: bacteria-like RNase P RNAs, mobile elements and a close source of the group I intron invasion in angiosperms. Nucleic Acids Research 33: 734–744.
31. Stockinger H, Walker C, Schüßler A (2009) 'Glomus intraradices DAOM197198', a model fungus in arbuscular mycorrhiza research, is not Glomus intraradices. New Phytologist 183: 1176–1187.
32. Krüger M, Krüger C, Walker C, Stockinger H, Schüßler A (2012) Phylogenetic reference data for systematics and phylotaxonomy of arbuscular mycorrhizal fungi from phylum to species level. New Phytologist 193: 970–984.
33. Pelin A, Pombert J-F, Salvioli A, Bonen L, Bonfante P, et al. (2012) The mitochondrial genome of the arbuscular mycorrhizal fungus Gigaspora margarita reveals two unsuspected trans-splicing events of group I introns. New Phytologist 194(3): 836–845.
34. Nadimi M, Beaudet D, Forget L, Hijri M, Lang BF (2012) Group I intron-mediated trans-splicing in mitochondria of Gigaspora rosea, and a robust phylogenetic affiliation of arbuscular mycorrhizal fungi with Mortierellales. Molecular Biology and Evolution. doi: 10.1093/molbev/mss088.
35. Hausner G (2012) Introns, Mobile Elements, and Plasmids. In: Bullerwell CE, editor. Organelle Genetics. Berlin: Springer. 329–357.
36. Griffiths AJF, Yang X (1995) Recombination between heterologous linear and circular mitochondrial plasmids in the fungus *Neurospora*. Molecular and General Genetics 249: 25–36.
37. Klassen R, Meinhardt F (2007) Linear protein primed replicating plasmids in eukaryotic microbes. Microbial linear plasmids. Berlin: Springer. 188–226.
38. Kennel JC, Cohen SM (2004) Fungal mitochondria: genomes, genetic elements and gene expression. In: Arora DK, editor. The handbook of fungal biotechnology 2nd ed. New York: Marcel Dekker Inc. 131–143.
39. Brown GG, Zhang M (1995) Mitochondrial plasmids: DNA and RNA. In: Levings CS III, Vasil IK, editors. The molecular biology of plant mitochondria. Dordrecht: Kluwer. 61–91.
40. Bertrand H, Griffiths AJF (1989) Linear plasmids that integrate into mitochondrial DNA in Neurospora. Genome 31: 155–159.
41. Cahan P, Kennell J (2005) Identification and distribution of sequences having similarity to mitochondrial plasmids in mitochondrial genomes of filamentous fungi. Molecular Genetics and Genomics 273: 462–473.
42. Ferandon C, Chatel SEK, Castandet B, Castroviejo M, Barroso G (2008) The Agrocybe aegerita mitochondrial genome contains two inverted repeats of the nad4 gene arisen by duplication on both sides of a linear plasmid integration site. Fungal Genetics and Biology 45: 292–301.
43. Dujon B (1980) Sequence of the intron and flanking exons of the mitochondrial 21S rRNA gene of yeast strains having different alleles at the co and rib-1 loci. Cell 20: 185–197.
44. Bell-Pedersen D, Quirk S, Clyman J, Belfort M (1990) Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications. Nucleic Acids Research 18: 3763–3770.
45. Dalgaard JZ, Garrett RA, Belfort M (1993) A site-specific endonuclease encoded by a typical archaeal intron. Proceedings of the National Academy of Sciences 90: 5414–5417.
46. Sharma M, Ellis RL, Hinton DM (1992) Identification of a family of bacteriophage T4 genes encoding proteins similar to those present in group I introns of fungi and phage. Proceedings of the National Academy of Sciences 89: 6658–6662.
47. Eddy S (1992) Introns in the T-seven bacteriophages [Dissertation]. Boulder: University of Colorado.
48. Paquin B, Laforest MJ, Lang BF (1994) Interspecific transfer of mitochondrial genes in fungi and creation of a homologous hybrid gene. Proceedings of the National Academy of Sciences 91: 11807–11810.
49. Bécard G, Fortin JA (1988) Early events of vesicular–arbuscular mycorrhiza formation on Ri T-DNA transformed roots. New Phytologist 108: 211–218.
50. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood,

Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution. doi: 10.1093/molbev/msr121.

51. Cho Y, Qiu Y-L, Kuhlman P, Palmer JD (1998) Explosive invasion of plant mitochondria by a group I intron. Proceedings of the National Academy of Sciences 95: 14244–14249.
52. Sanchez-Puerta MV, Cho Y, Mower JP, Alverson AJ, Palmer JD (2008) Frequent, Phylogenetically Local Horizontal Transfer of the cox1 Group I Intron in Flowering Plant Mitochondria. Molecular Biology and Evolution 25: 1762–1777.
53. Robison MM, Wolyn DJ (2005) A mitochondrial plasmid and plasmid-like RNA and DNA polymerases encoded within the mitochondrial genome of carrot *Daucus carota*. Current Genetics 47: 57–66.
54. Brügger K, Torarinsson E, Redder P, Chen L, Garrett R (2004) Shuffling of Sulfolobus genomes by autonomous and non-autonomous mobile elements. Biochemical Society Transactions 32: 179–183.
55. Biessmann H, Valgeirsdottir K, Lofsky A, Chin C, Ginther B, et al. (1992) HeT-A, a transposable element specifically involved in "healing" broken chromosome ends in Drosophila melanogaster. Molecular and Cellular Biology 12: 3910–3918.
56. Fricova D, Valach M, Farkas Z, Pfeiffer I, Kucsera J, et al. (2010) The mitochondrial genome of the pathogenic yeast Candida subhashii: GC-rich linear DNA with a protein covalently attached to the 5′ termini. Microbiology 156: 2153–2163.
57. Foury F, Roganti T, Lecrenier N, Purnelle B (1998) The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae*. FEBS Lett 440: 325–331.