Year : 2017

Genetic diversity of the symbiotic fungus: Rhizophagus irregularis and its effect on a plant host and a plant community

Savary Romain

Savary Romain, 2017, Genetic diversity of the symbiotic fungus: Rhizophagus irregularis and its effect on a plant host and a plant community

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive http://serval.unil.ch
Document URN : urn:nbn:ch:serval-BIB_E8D4F131E0266

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Department of Ecology and Evolution

Genetic diversity of the symbiotic fungus: *Rhizophagus irregularis* and its effect on a plant host and a plant community

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de Biologie et de médecine
de l'Université de Lausanne

par

Romain Savary

Biologiste diplômé d'un MSc en Comportement, Évolution et Conservation
Université de Lausanne, Suisse

Jury

Prof. Medhi Tafti, Président
Prof. Ian R. Sanders, Directeur de thèse
Prof. Nicolas Salamin, Expert interne
Prof. Bruce McDonald, Expert externe

Lausanne - 2017

to my family

Table of Contents

Acknowledgements

I spent my entire undergraduate studies as well as this PhD at the University of Lausanne, thus the DEE and all the people that passed in Lausanne contributed significantly to an amazing part of my life.

First of all I would like to thanks Prof. Ian R. Sanders for having me offered the opportunity to accomplish my PhD in his group. I would like also to thanks Ian for all he taught me, his trust, his encouragement, all the freedom he offered for research and his humanity in believing that research can bring something bigger than only science. Working with Ian was a real pleasure.
A particular acknowledgment to my experts Prof. Nicolas Salamin and Prof. Bruce McDonald for having accepted to read and comment on this work and for their precious time. I also acknowledge the president of the jury, Prof. Medhi Tafti.

I particularly thank Frédéric Masclaux, Tania Wyss Lozano Hoyos, Christophe Dufresnes, Lucas Villard and Ivan Mateus Gonzales for the things I have learnt by working by their side as well as all past and present members of the Sanders groups. Among them a particular thanks to Jérôme Wassef, Lucas Villard, Cindy Dupuis, Jérémy Bonvin, Cristian Camillo Ricon, Joaquim Cruz Corella, William Chanz Robbins, Ricardo Pena Venegas and Alia Rodriguez. I would also like to thank all the students that helped me to produce this work, particularly my Master student, Cindy Dupuis as well as other students, Laura Alvarez Lorenzana, Marie Strehler, Cynthia Meizoso, Noémie Gambino, Nicolo Tartini, Rafael Joss, Ana Paula Machado and all Prof., particularly Prof. Nicolas Perrin as well as all assistants and students from the Roscoff internship. Thanks to all the technicians for their priceless help, Nicolas Ruch, Jérémy Bonvin, Consolée Alleti, Catherine Berney and Roberto Sermier.

There are so many other people to thank for their friendship and their love that helped me to achieve this thesis especially people from the penthouse and the mansion. I will not cite them all here in order not to forget anyone but I think that these people will recognize themselves. However I want to specially thank my parents, Edouard and Fabienne Savary, my sister, Céline Savary and my uncle and godfather, Jérôme Meizoz for their support and finally Athimed El Taher for her love and her presence.

Abstract

Arbuscular mycorrhizal fungi (AMF) are endosymbionts of the vast majority of terrestrial plants species exchanging soil nutrient for plant carbohydrates. This symbiotic association evolved more than 600 Myrs ago and might have helped the colonization of land by plants. AMF have been shown to confer a large number of benefits to plants, such as resistance to drought, pathogens, salt and pollutants. They were also shown to be central in terrestrial ecosystems as AMF diversity was found to impact plant community diversity, structure and productivity. Up to now, AMF diversity was measured to around 348 virtual molecular taxa for approximately 200'000 host plants. It is surprising that during their long evolution, the AMF did not radiate in species number, as did their hosts. In the first part of the thesis, I thus investigated the possibility that AMF could hide more functional diversity than previously thought. A diversity that could have been hidden until now because traditional molecular methods used could have had a to low resolution. Therefore using double digested restriction-site associated DNA sequencing (ddRAD-seq) protocols on the model AMF, *Rhizophagus irregularis*, I found that ecologists probably underestimated the functional diversity of AMF. Indeed I found well-defined genetic groups within this model species with differential transcriptome expression and differences in phenotypic traits. I also confirmed the controversial results suggesting a low endemism of AMF. In the second part, I then tested the hypothesis that within species AMF diversity could be ecologically functional. The intra-specific diversity was then inoculated either as a single isolate inoculum or as a mix inoculum on simulated plant communities. I found that phylogenetically similar AMF tend to impact the plant community in the same way by more or less repressing the dominant plant of the community, thus, resulting in a change in resource partitioning among subordinate plants. In the third part I characterized the molecular interaction between within species functional AMF diversity and the globally important crop plant *Manihot esculenta*. By performing a dual-transcriptome sequencing experiment I was able to unravel important plant genes and fungal genes that could partly explain how the AM symbiosis could switch from one extreme to another along the symbiosis continuum, parasitic to mutualist. Finally I have also found that *R. irregularis* hosts a rare and probably parasitic endosymbotic bacterium adapted to the intracytoplasmic life. This thesis showed several plant and fungal factors explaining the variability in the outcome of this ecologically and agriculturally important symbiosis.

Résumé

Les champignons endosymbiotiques à arbuscules (CEA) forment une symbiose avec la vaste majorité des plantes terrestres en échangeant des nutriments du sol contre des sucres issus de la photosynthèse. Cette association à évolué il y a plus de 600 millions d'années et a pu être d'une grande aide pour la colonisation du milieu terrestre par les plantes. Il a été démontré que la symbiose avec CEA peut apporter de nombreux bienfait aux plantes, comme une meilleure résistance à la sécheresse, aux pathogènes, au sel ainsi qu'aux polluants. De plus, les CEA s'avèrent jouer un rôle central dans les écosystèmes de plantes de par le fait que la diversité des CEA influence leur diversité, leur structure ainsi que leur productivité. De nos jours, la diversité des CEA est évaluée à 348 espèces moléculaires, pour environ 200'000 plantes hôtes. Il est surprenant cependant, après une évolution aussi longue, que si peu d'espèces soient trouvées chez les CEA comparées à leurs plantes hôtes. La première partie de cette thèse m'a permis de tester l'hypothèse qu'il existe une diversité demeurée cachée chez les CEA, du fait que les méthodes moléculaires utilisées jusque là avaient une trop faible résolution. A travers l'utilisation d'un protocole ddRAD-seq sur l'espèce modèle des CEA, *Rhizophagus irregularis*, j'ai pu déterminer que les écologistes moléculaires sous-estiment très probablement leur diversité. En effet, j'ai identifié quatre groupes génétiques bien définis avec de fortes différences transcriptomiques ainsi que phénotypiques. Dans la foulée, ce travail confirme les résultats débattus sur le faible taux d'endémisme trouvé chez les CEA. Dans la deuxième partie de cette thèse j'ai voulu savoir si cette nouvelle diversité pouvait jouer un rôle écologiquement fonctionnel. J'ai donc inoculé cette diversité, soit sous la forme d'une seule souche soit sous la forme d'un mix de souches, sur des communautés de plantes simulées en serre. Il en résulte que des CEA phylogénétiquement similaires tendent à influencer dans le même sens les communautés de plantes en réprimant, plus au moins selon les clades, la dominance d'une plante et permettant ainsi un changement dans la répartition des ressources entre les plantes secondaires. Dans la troisième partie de cette thèse, je me suis intéressé à l'effet de cette diversité sur la réponse phénotypique et moléculaire d'une espèce de végétale, le manioc (*Manihot esculenta*), une plante agricole de grande importance alimentaire. Par le biais d'une expérience de double séquençage d'ARN, j'ai pu identifier des gènes de plantes et de champignons importants pour la symbiose qui pourraient expliquer comment cette symbiose passe de parasitique à mutualistique. Finalement, j'ai identifié une bactérie endosymbiotique complètement adaptée à la vie intracytoplasmique dans *Rhizophagus irregularis*. Cette thèse apporte de nombreux éléments permettant de mieux comprendre la variété d'interactions existant entre les CEA et les plantes ainsi que leurs communautés.

Chapter 1

General introduction

Diversity of living forms

All life is founded on the principle of diversity or variability. Diversity of forms and of behavior, giving rise to astonishing shape from the tiny body of the tree frog made for climbing to the huge and hydrodynamic shape of the blue whale. This diversity of living forms arose from the diversity of natural habitats available on earth and by the competition among organisms for obligatory resources necessary to sustain life. Every niche with available resources were occupied by the best competitors, that slowly changed their phenotypes to be more and more adapted and specialized for resource collection. Some organisms even specialized on other organisms as a food source. Their behavior ranging from completely harmless mutualisms where both organisms profit to complete pathogens that severely reduces the capability of the host to survive and reproduce. All such biotic interactions are described under the general term, symbiosis. Symbiosis was argued to be a strong evolutive force, allowing some of the major transitions of life on earth (O'Malley et *al.,* 2014). The dynamics of these symbiotic relations is often described as an arms race following the Red queen hypothesis (Van Valen, 1973), where both the host and the symbiont try to improve its' weapons to escape or control the other. Arms races can only occur if the host and the symbiont show diversity in the efficiency of the tools or weapons used to achieve the symbiosis. Such variability will permit a dynamic equilibrium, where occasionally individual in the population will lose the arms race for example because of a weaker defense system while some other individuals will escape the control of the symbiont. Such dynamic equilibrium could in some cases lead to a co-evolution of the host and the symbiont, where both are highly specialized and dependent on each other.

One of these arms races is the arms race between plants and fungi. The kingdom of fungi has developed a large variety of strategies to obtain nutrients from the plant, from pathogenic to mutualistic (James et *al.,* 2006). In this thesis I have focused on probably the most common, and possibly the longest arms race between two terrestrial organisms, between arbuscular mycorrhizal fungi (AMF) and land plants.

Arbuscular mycorrhizal fungi (AMF) symbiosis

Arbuscular mycorrhizal fungi (AMF) are the major symbionts of terrestrial plants ecosystems from polar tundra to tropical forests, forming symbiotic interactions with more than 74% of land plants species (van der Heijden et *al.*, 2015). Each partner was shown to profit from this symbiosis. The gain of the plant in this symbiosis is mainly found in an increased amount of the nutrients taken up, such as phosphate (Karandashov & Bucher, 2005) and nitrate (Govindarajulu et *al.*, 2005) that is provided by the so-called fungal "extended root system". The fungus conversely, obtains carbon from the plant in the form of photosynthetic sugar (Bago et *al.*, 2003). These fungi have specialized endocellular structures called the arbuscules (Fig. 1). These tiny tree-like structures found within the plant root cells are the center where exchange between the plant and the fungus occur (Parniske, 2008). AMF are composed of three other main structures. First, the hypha that comprises the mycelial network and that collects nutrients is found inside the root as the intra-radical mycelium (IRM) or outside the roots as the extra-radical mycelium (ERM). The hyphae are continuous such as no septa are formed such that, the cytoplasm, the multiple nuclei coexisting within the cytoplasm and all organelles and nutrients are potentially free to move from one extremity to another. Second, the spores are the reproductive organs representing propagules for dispersion. Spores are produced in large numbers and are filled with the flowing cytoplasm with large number of nuclei and organelles (Fig. 1). Spores are vegetative reproductive organs and no sexual reproduction was ever recorded in AMF, despite several clues indicating recombination (Vandenkoornhuyse et *al.*, 2001, Croll et *al.*, 2009, Ropars et *al.*, 2016) as well as the findings of conserved meiotic genes (Halary et *al.*, 2011). Thus, AMF have been characterized as ancient asexuals. Third, the vesicles are the storage organs of the fungi and are usually found in the roots of the plant (Fig. 1).

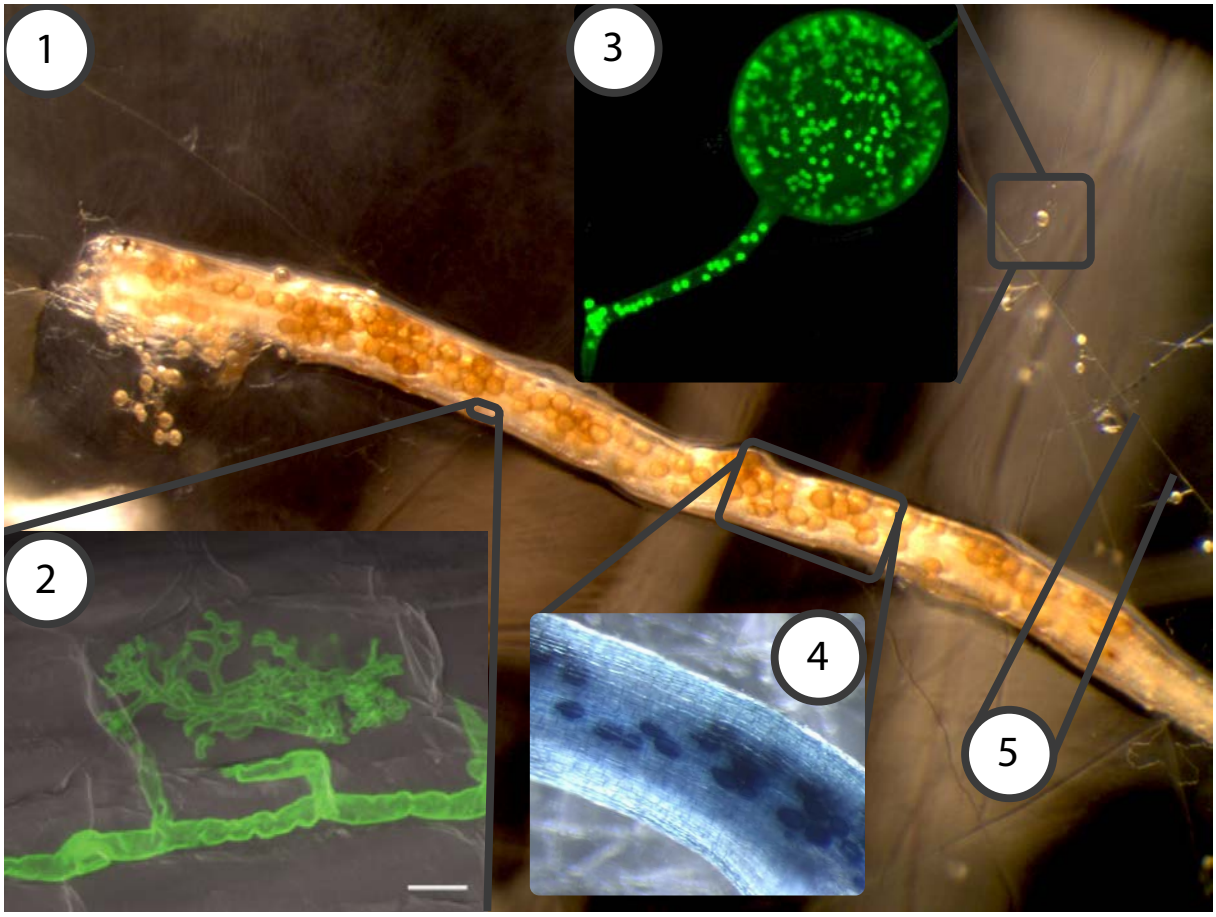


Fig.1 (1) Roots colonized by an AMF (*Rhizophagus irregularis*, © J. Bonvin) (2) Arbuscule within a plant root cell (Bravo et *al.*, 2017) (3) An AMF spore with multiple nuclei (© M. Hijri) (4) Vesicules within a root (5) extra radical hyphae.

Evolution of the AMF

Fossils of probable AMF were found with an age of up to 460 million years and helped the molecular dating to date the evolution of this symbiosis around 600 million years ago (Redecker et *al.*, 2000). With such an early appearance of these fungi, it has been suggested that they could have potentially played a central role in the colonization of land by plants (Heckman et *al.*, 2001, Parniske, 2008,). Despite having evolved more than 600 millions years ago, and having helped plants to make an important transition, it is surprising that the diversity of AMF has not flourished as has the diversity of plants. AMF diversity is restricted to only 300 to 1600 described morphological or molecularly characterized species. This low diversity could be the results of one of the keys of their success; their generalist life style with low host specificity, allowing them to colonize many plants. This faculty is somehow necessary for such obligatory fungi that could not survive without colonizing a plant host. The

obligatory biotrophic lifestyle of these organisms could suggest that AMF are highly specialized on some plant species or group of species. All efforts to find co-evolution of individual AMF species with individual plant species have been unsuccessful. The AMF taxonomic group remains as a group of species with low endemism according to the recent findinds of Davison et *al.*, in 2016. Despite these findings of global low endemism of AMF, suggesting a low restriction of AMF species distribution to biomes, plant families or plant species, it has been shown that different plant species can harbour different AMF communities (Vandenkoornhuyse et *al.*, 2002 & 2003). Moreover, at the within species level it has even been suggested that certain genotypes of *R. irregularis* could prefer one plant or another during the trap culture process used to isolate AMF from the field (Croll et *al.*, 2008). These experiemental results contradict the findings of Davison et *al.*, 2016 and support a specialization of the AMF species and genotypes to different plants.

Even if AMF are obligatory biotrophs, they are not necessary strong competitive excluders. Thus it is common to find more than one species of AMF on a plant (Maherali & Klironomos, 2007). It is also common for an AMF to colonize more than one plant at a time (Kiers et *al.*, 2011). All these interactions, thus, suggest highly complex networks where all the members of a community of AMF interact with a large number of individual plants in the community.
Due to the complex numbers of combinations of AMF species, plant species and types of soil, the outcome of the symbiosis is plastic and can range from a mutualistic to a parasitic interaction (Klironomos, 2003). However, the mean outcome of the AMF-plant symbiosis is highly likely to be mutualistic. As a testimony we can cite the ~600 millions year of persistence of the interaction and the development of a specialized plant gene toolkit for AMF interactions, with a set of genes that are only conserved in the plants that engage into AMF symbiosis (Delaux et *al.*, 2014, Bravo et *al.*, 2016). The outcome of this symbiosis on the long term has to be positive on the fitness on both partners to observe such wide spread ecological success.

Communities, species, populations and individuals

The definitions of species and all deriving concepts of community, population and individuals commonly used in most vertebrates, invertebrates or plants should be readapted to AMF (Rosendahl 2008). Indeed the definition does not fit these organisms that lack a known sexual

reproduction and have such particular intra-cellular organisation including a free cytoplasmic population of nuclei suggested to be transferred from AMF to AMF by the process of anastomosis i. e. the fusion of hyphae.

For a long time morphological characteristics of AMF spores was the way to describe an AMF species. The development of molecular methods has boosted the discovery of new AMF species by using common barcode markers, mostly ITS and the small and large subunits of the ribosomal genes (SSU and LSU, Krüger et *al.,* 2012). The rise of the early molecular methods came with the development of the molecular definition of a species in AMF. The molecular species are described under the term of operational taxonomic unit (OTU) or virtual taxa (VT) and are based on arbitrary limits with a certain threshold of divergence between sequences (Opik et *al.,* 2010). However, as the reproductive mode is not clearly stated, it is hard to define the boundary between two species where the gene flow has ceased. The mechanism of anastomosis between two different AMF was shown to be a way for AMF to exchange genetic information (Croll et *al.,* 2009). Moreover, ability of performing perfect fusion between two isolates was shown to depend on the genetic distance. Indeed the meeting of the hyphae of two time the same clonal individual resulted in ~50% of perfect fusion, by increasing the genetic Euclidian distance between the two meeting AMF, suddenly these values dropped drastically and quickly to 5, 1 and finally 0%. A way to define the boundary of AMF species would then be to find the optimal genetic distance at which anastomosis no longer results in the persistence of the exchanged nuclei across generations. However, nothing is known about such persistence across generations.

The definition of an individual in counterpart has been made clear, once AMF started to be isolated in *in vitro* conditions. In AMF, an individual is one spore that when through all the process of isolation. This process is going from the natural soil, to single spores culture and finally into *in vitro* culture with Ri-T DNA carrot roots, such that a new culture started from one single spore is commonly described as an isolate. This isolate will then be conserved, shared between laboratories and archived in AMF bank. As the process of isolation is difficult to achieve, highly time consuming (1-2 years) and does not work well for all AMF species (Koch et *al.,* 2004), most research to date has focused on the model *R. irregularis* isolate DAOM197198 and on very few species of mainly the genus *Rhizophagus*.

All isolates of this work are individuals that went through several generations of culturing and each isolate originated from a single spore that was cultured in order to maintain the original genetic makeup.

Genetics and Genomics of AMF

With the first publication of an AMF genome in 2013, the genome of *Rhizophagus irregularis* (isolate DAOM197198, Tisserant et *al.,* 2013) and later on the publications of 4 single nucleus genomes of the same *R. irregularis* isolate (Lin et *al.,* 2014), 5 *R. irregularis* genomes (Ropars et *al.,* 2016) and ddRADseq data on a population of 20 *R. irregularis* isolates (Wyss et *al.,* 2016) great improvements were made in understanding the nuclear organisation and the molecular mechanisms of this fungal symbiont.

In parallel to this deep characterization of one species of AMF, similar sequencing technologies were used for AMF community profiling. Such analysis performed via the amplification and the sequencing of a single marker of around 500bp, from DNA of a soil or root sample, resulted often in hundred of thousands of sequences per sample, allowing the characterization of AM fungal communities of an ecosystem, of a plant species or a plant family. Up to now, the most important findings that came out of a study of worldwide AMF community profiling showed no continental partitioning as 93% of the taxa were found at least on two continents and 34% of the taxa were found on all continents, leading to the conclusion that AMF is a group with low worldwide endemism (Davison et *al.,* 2016). Such surprising findings immediately raised criticism. Bruns and Taylor (2016) advanced the argument that if mammals were analysed with the 500bp portion of the same maker, and clustered at the same level as the sequences used for AMF profiling, we would classify all the mammals in the same species and this species would then be widely spread on all continents.

The gap between the deep characterization of few isolates and the low molecular resolution of the AMF diversity by community profiling needs to be filled. In **chapter 2** we, thus, attempted to fill this gap by analysing large-scale diversity of *R. irregularis* across continents with deep sequencing methods.

Ecological role of AMF

As said earlier, AM fungal symbiosis is the most frequent terrestrial symbiosis involving plants and, thus, it could be central in ecological processes. Indeed, it was shown that up to 20% of the carbon fixed by the plants is transferred to the fungi, demonstrating their important role the carbon cycling (Parniske, 2008). More than being at the crossroad of

ecological processes as a passive actor, AMF were shown to act on the diversity of plant species such that the diversity of AMF contributed positively to the diversity of plant species (van der Heijden et *al.,* 1998). In parallel, AMF diversity was shown to increase the global productivity of plant communities and was also found to be important for plant community structure, as some isolate were found to increase the growth of specific plant while some others increased other plant species. In this context, isolates were also observed to reduce the growth of dominant plants within a community thereby allowing the increased growth of subordinate plants (Wagg et *al.,* 2011, Mariotte et *al.,* 2013). Several other benefits of the AMF symbiosis were observed on single plant species and could possibility be generalized to plant communities in the future. We can count among them the increase of inoculated plant resistance to drought, salinity stress, deprivation of nutrients, atrazine and other contaminants, pathogens and insects (Gonzalez-Chavez et *al.,* 2002, Al-Karaki et *al.,* 2004, Pozo & Azcón-Aguilar 2007, Koricheva et *al.,* 2009, Hajiboland et *al.,* 2010, Song et *al.,* 2015, Garcia et *al.,* 2017).

Intra-specific diversity of AMF

Following the findings of the impact of AMF diversity on plant diversity (van der Heijden et *al.,* 1998), the ecological role of AMF diversity was then strongly investigated. To our knowledge almost all ecological experiments used only one AMF isolate per species making the assumption that, first the species concept in AMF is a reliable concept and second that within an AMF species the effect on plant and plant communities is homogenous (van der Heijden et *al.,* 1998, Klironomos et *al.,* 2000, Maherali & Klironomos, 2007). However, boundaries defining AMF species are still not clear and even molecular data of small variable markers may not be precise enough to detect variants of different species. Moreover, Sanders & Rodriguez (2016) suggested that the second assumption on homogeny of the effect of intra AMF specific diversity on plant communities might be violated and that a more relevant and ecological functional diversity could be found at the intra-specific level in AMF. This hypothesis emerged due to the results of several studies (Munkvold et *al.,* 2004, Mensah et *al.,* 2015) where separately several isolates of the same AMF species were inoculated on one plant species. This process was then repeated for other species of AMF. These studies concluded that the variability in plant growth response such as root, shoot dry mass and phosphate uptake and in fungal traits could be higher within AMF species than among

isolates of different species. A third study (Koch et *al.,* 2006) found similar results of high variability in plant response but within only one population of *R. irregularis* isolates.
From these results Sanders & Rodriguez (2016) suggested that intra-isolate diversity should be investigated as a potential driver of plant diversity and plant community structure. It would then be important to have a well-defined molecular diversity for several species of AMF. In **chapter 2** we made a step in that direction by characterizing deeply the diversity of *R. irregularis* for a large number of isolates of *R. irregularis* from different geographical origins. The information in chapter 2 on the genetic divergence among isolates was then used in **chapter 3** to investigate the effect of this within-AMF species genetic diversity as a potential source of functional variation impacting the stability of a simulated calcareous grassland plant community.

Agronomic role of AMF

With their low host specificity, the variability in plant response to different AMF and their capability in poor soil fertility to acquire and transmit phosphorus to plants, AMF were considered as ideal candidates for agronomic program to increase the yield of important crop plants (Rodriguez & Sanders, 2015). AMF can perform definitely symbiosis with the major crops plants and have the capacity to increase their growth, such as wheat, rice and cassava (Pellegrino et *al.,* in 2015, Angelard et *al.,* 2010 and Ceballos et *al.,* 2013). The production of *Manihot esculenta* (cassava) was recently shown to be improved by the inoculation with *R. irregularis* in real farming conditions (Ceballos et *al.,* 2013), this study paved the way for the use of AMF in agriculture, especially in order to increase important food security crops in tropical countries such as the cassava. Indeed, tropical soils are ideal for the symbiosis with AMF as they are deprived of phosphate.
The functional diversity suggested within AMF species, and found as the result of the inoculation of several isolates of the same species on several plants, might also be true for crops plant species. In order to use these fungi in the future for a more productive and sustainable agriculture, a better understanding of the effect of this within-AMF species variability on crop plant is needed.
The within-species diversity that we described in *R. irregularis* in **chapter 2** was, thus, used in **chapter 4** in a greenhouse experiment, where we inoculated clonal cassava plants with 12 *R. irregularis* isolates, evenly spread across the phylogenetic diversity of this species, in order

to investigate the impact of this potential within-species diversity on an important crop plant. The impact of the diversity of *R. irregularis* isolates was investigated in terms of phenotypic measurements as well as transcriptome patterns. In this chapter, a dual-RNAseq was performed to understand at the molecular level the interaction of the expression of cassava genes and *R. irregularis* genes from different isolates.

Authors contributions

Chapter 2: RS and IRS designed the study, RS collected, isolated, photographed, extracted and ddRADseq sequenced all isolates with the help of TW. JBM gave some isolates, RS analysed the ddRADseq with the help of FGM. GD isolated the African fungi from the Ivory Coast, RS sequenced the phosphate transporter gene. RS cloned and sequenced the ITS fragment with the help of JCC. RS and APM measured phenotypic traits. RS analysed all the data. RS drafted the manuscript with the help of IRS and JBM

Chapter 3: RS and LV conceived the ideas and designed the study. RS and LV conducted the experiment and collected the data. RS and LV analyzed the data. All authors contributed critically to the writing of the manuscript.

Chapter 4: RS and IMG designed the study with the help of CD, ECRT and IRS. RS, IMG performed the fungal and plant *in vitro* culturing. RS, IMG, CD and ECRT set up the experiment and harvested the experiment. RS and CD performed the phenotypic measurements, the RNA extractions, constructed and sequenced the RNA-seq libraries. RS analysed all the data with the help of FGM. RS wrote the manuscript with IRS.

Chapter 5: RS designed the study, sequenced and analysed all the data. RS sequenced the genome of the bacteria and the fungi and is analysing the data with FGM. RS drafted the manuscript with the help of IRS.

Chapter 2

A population genomics approach shows widespread geographical distribution of cryptic genomic forms of the symbiotic fungus *Rhizophagus irregularis*

Romain Savary[1], Frédéric G. Masclaux[1, 2], Tania Wyss[1], Germain Droh[1, 3], Joaquim Cruz Corella[1], Ana Paula Machado[1], Joseph B. Morton[4] and Ian R. Sanders[1].

[1] Department of Ecology and Evolution, Biophore Building, University of Lausanne, 1015 Lausanne, Switzerland. [2] Vital-IT Group, SIB Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. [3] Laboratoire de Génétique, Unité de Formation et de Recherche en Biosciences, Université Félix Houphouet Boigny, 22 BP 582 Abidjan 22, Ivory Coast. [4] Division of Plant and Soil Sciences, West Virginia University, Morgantown, WV 26506, USA.

This chapter was accepted for publication in the ISME journal in August 21, 2017

Abstract

Arbuscular mycorrhizal fungi (AMF; phylum Gomeromycota) associate with plants forming one of the most successful microbe-plant associations. The fungi promote plant diversity and have a potentially important role in global agriculture. Plant growth depends on both inter- and intra-specific variation in AMF. It was recently reported that an unusually large number of AMF taxa have an intercontinental distribution, suggesting long-distance gene flow for many AMF species, facilitated by either long-distance natural dispersal mechanisms or human-assisted dispersal. However, the intercontinental distribution of AMF species has been questioned because the use of very low resolution markers may be unsuitable to detect genetic differences among geographically separated AMF, as seen with some other fungi. This has been untestable because of the lack of population genomic data, with high resolution, for any AMF taxa. Here we use phylogenetics and population genomics to test for intraspecific variation in *Rhizophagus irregularis,* an AMF species for which genome sequence information already exists. We used ddRAD-sequencing to obtain thousands of markers distributed across the genomes of 81 *R. irregularis* isolates and related species. Based on 6 888 variable positions, we observed significant genetic divergence into 4 main genetic groups within *R. irregularis,* highlighting that previous studies have not captured underlying genetic variation. Despite considerable genetic divergence, surprisingly, the variation could not be explained by geographical origin, thus also supporting the hypothesis for at least one AMF species of widely dispersed AMF genotypes at an intercontinental scale. Such information is crucial for understanding AMF ecology, and how these fungi can be used in an environmentally safe way in distant locations.

Key words: ddRAD-sequencing, AMF, *Rhizophagus irregularis*, Population genomics, Symbiosis, microorganisms.

Introduction

In 1934, Baas-Becking was the first to propose a worldwide distribution of microorganisms, saying that 'Everything is everywhere, the environment selects'. A recent study of the global distribution of arbuscular mycorrhizal fungi (AMF), comprising the phylum Glomeromycota, suggests that the statement by Baas-Becking might be true for this phylum. Using virtual taxa (VT; Öpik *et al.*, 2010), discriminated by a 520 bp region of the 18S (small sub-unit or SSU) rRNA gene, 93% of all AMF taxa were found to be present on more than one continent and 34% of species occurred on all continents except Antarctica (Davison *et al.*, 2015). Global distribution of one-third of a phylum in eukaryotes is particularly unusual, especially in organisms with a low dispersal rates (Tedersoo *et al.*, 2014). Records of cosmopolitan species, excluding marine (Finlay, 2002) and invasive or pest species (Margaritopoulos *et al.*, 2009), are rare. Birds have great dispersal capabilities, but even so, only six out of 10 000 species have a cosmopolitan distribution. Even those six species exhibit distinct genetic substructure (Monti *et al.*, 2015).

Several findings in fungal population genetics suggest that, for some species, a cosmopolitan distribution reflects hidden endemism (Taylor *et al.*, 2006). Some species first thought to be cosmopolitan are actually comprised of cryptic species with distinct geographical ranges such as the Basidiomycete *Schizophyllum commune* (James et *al.*, 1999), *Aspergilllus fumigatus* and six other cited Ascomycete species (Pringle *et al.*, 2005).

Morton (1990) hypothesized that AMF speciation occurred prior to the breakup of Pangea and that fungal lineages co-evolved with the emergence of plant species thereafter, thus, resulting in global distribution. Alternatively, Davison *et al.*, (2015) favored a hypothesis based on a more recent VT-molecular clock phylogeny that such widespread distribution was the result of a recent dispersal of spores, mediated by strong storms, as well as previously underestimated human activities. In either case, interpretations may be misleading if there is an underestimation of true species and an overestimation of species with an intercontinental distribution. The use of such short SSU sequences and a 97% similarity threshold to construct VT taxa (Davison *et al.*, 2015) might not be informative enough to resolve significant AMF clades (Bruns and Taylor, 2016; Schlaeppi *et al.*, 2016). Their argument is based on the premise that two taxa that share a 97% similarity in their SSUs may have diverged genetically many millions of years ago into different species. Davison *et al.*, 2015 justified the use of VT because they are "phylogenetically defined sequence groups that exhibit a taxonomic

resolution similar to that of morphological species". However, the absence of clear morphological distinctions and indistinguishable life history traits between *R. intraradices* and *R. irregularis*, together with highly supported divergence based on rDNA sequences (Stockinger *et al.*, 2009), provides compelling evidence of cryptic speciation in the *Rhizophagus* clade. Similarly, other species may be genetically divergent but are morphologically similar (Rosendahl *et al.*, 2009). This is more likely to occur in small eukaryotes with limited phenotypic space for evolution of new traits but with a less constrained capacity for genotypic evolution (Taylor *et al.*, 2000). Discovery of these cryptic taxa requires the sampling of globally distributed AMF taxa and the application of population genomics tools because a large number of markers distributed across the genome would give a finer resolution suitable to detect genetic differences among populations (Bruns and Taylor, 2016; Öpik *et al.*, 2016).

A high throughput sequencing technique, double digested restriction-site associated DNA sequencing (ddRAD-seq) is a reliable way of obtaining a large number of genome-wide markers (Parchman *et al.*, 2012; Peterson *et al.*, 2012). This overcomes bias created by use of single gene sequences and low similarity thresholds (e.g., the SSU rRNA gene). This approach was previously applied to *Rhizophagus irregularis*, an AMF species that has a putative global distribution and can be cultured under *in vitro* conditions. *In vitro* culturing allows the extraction of DNA that is free of contaminant microorganisms or plant DNA. This technique was previously shown to be reliable for *Fusarium* (Talas and McDonald, 2015) and in a study measuring inter- and intra- isolate variation in a population of *R. irregularis* (Wyss *et al.*, 2016). This multi-locus approach can provide a much finer resolution of the divergence between different *R. irregularis* isolates, which in turn can provide data to resolve whether this species has evolved under strong or relaxed geographical constraints.

Describing the genetic diversity of AMF species, and especially the genetic diversity of *R. irregularis* at a wide geographical scale, is fundamental to understanding biogeographic patterns. However, it is not only important for resolving biogeographic questions but also has strong implications for interpretation of ecological, agronomic, and environmental studies. First, AMF are important symbiotic partners that interact with many plant species. Approximately 300 – 1 600 predicted AMF taxa form symbioses with over 200 000 plant species (van der Heijden *et al.*, 2015). The fungi have the capacity to efficiently absorb nutrients, particularly phosphate, from the soil and give them to the plant (Smith and Read, 2010) and richness of AMF taxa promotes plant diversity (van der Heijden *et al.*, 1998). Most

ecological studies have not considered the role of intra-specific AMF variation even though intra-specific differences in AMF can have larger effects on P uptake and variation in plant growth than inter-specific differences (Munkvold *et al.*, 2004; Mensah *et al.*, 2015; Rodriguez and Sanders, 2015). Thus, measurements of genetic variation within AMF species could be highly ecologically relevant. Second, genetic variability in *R. irregularis* population causes significant variation in plant growth (Koch *et al.*, 2006). The potential of using naturally existing genetic variation in this species to develop more efficient strains to increase plant growth has been demonstrated by Angelard *et al.*, (2010) where five-fold differences in rice growth could be achieved due to genetic variation in *R. irregularis* strains. Defining the total genetic variation in this AMF species is important for future programs using the genetic variation in this fungus for more efficient inoculants (Sanders, 2010). Third, *in vitro* grown *R. irregularis* significantly increases yields of the globally important crop cassava (Ceballos *et al.*, 2013). *R. irregularis* is potentially a good candidate species for large-scale inoculation of tropical crops because it can be mass-produced in contaminant-free *in vitro* conditions and because it appears to have a global distribution. However, before introducing high growth-promoting *R. irregularis* isolates from one location to another, it is important to establish with population genomics techniques the risk of introducing genetically novel isolates into a new environment with the potential to become invasive. If *R. irregularis* has a very wide geographical distribution of very genetically similar isolates then the risk of introducing exotic genetic material is low. If there is strong geographically determined genetic structure among *R. irregularis* populations then such a risk is much higher (Rodriguez and Sanders, 2015).

In this study, 61 isolates identified morphologically and genetically as *R. irregularis* were obtained from various locations across Europe, North Africa, Middle East and North America and were propagated *in vitro* and sequenced with ddRAD-seq. To the 61 isolates, data from 20 other isolates, that had been sequenced previously with the same ddRAD-seq protocol by Wyss et *al.* (2016), were added. We addressed three central questions: (i) Is the low level of endemism reported by Davison *et al.*, (2015) applicable to *R. irregularis* and related species when using high resolution genome wide markers that are capable of showing within population genetic variation? (ii) Is there significant genetic diversification among continents? (iii) Is there evidence that some quantitative traits vary in accordance with the genetic variation observed in this fungal species?

Material & Methods

Fungal isolates and culturing

Rhizophagus irregularis (known previously as *Glomus intraradices, Glomus irregulare* or more recently as *Rhizoglomus irregulare*) were obtained from collections and biobanks such as BEG (http://www.i-beg.eu), GINCO (http://www.mycorrhiza.be/ginco-bel/) and INVAM (http://www.invam.wvu.edu), from small enterprises: Symbiom (Czech Republic; https://www.symbiom.cz/en/ and INOQ (Germany), and from research groups (IRTA: http://www.irta.cat, TERI: http://www.teriin.org, or were present in our group at the University of Lausanne. For more details on nomenclature, see Supplementary note S1. Sixty-one isolates were collected (Supplementary Table S1). The collection included the *R. irregularis* holotype, DAOM197198, cultured *in vitro* in two laboratories (Switzerland: DAOM197198-CH and Czech Republic: DAOM197198-CZ), the holotype of *R. intraradices*, FL208 grown *in vitro* from pot culture material provided by INVAM and *R. proliferus* that was used as an outgroup for some analyses.. All isolates in this study were cultured *in vitro* with Ri T-DNA transformed carrot roots in a two-compartment culture system (St-Arnaud *et al.*, 1996) to obtain fungal spores free from host plant DNA or other microorganisms. Some isolates were received as soil samples and were transferred to *in vitro* single spore cultures (Supporting information, Supplementary note S2).

DNA extraction, amplification and sequencing of the phosphate transporter and the SSU

Three-month-old sporulating *in vitro* cultures were used for DNA extraction. The medium in compartments containing only spores and hyphae was dissolved in 500ml of citrate buffer (0.0062 M of citric acid anhydrous and 0.0028 M of sodium citrate tribasic dihydrate) for 1h with a magnetic stirrer. Spores and hyphae were collected by filtering in citrate buffer through a sieve with 32µm openings and rinsed with ddH_20. Spores and hyphae were frozen in liquid nitrogen and ground with a sterile pestle. DNA was extracted with the DNeasy Plant Mini Kit (QIAGEN) following the manufacturer's instructions. Species identity of each isolate was verified by sequencing the Phosphate Transporter Gene (PTG), which is considered a phylogenetically informative marker by Sokolski *et al.*, (2011) .

The primers AML2 and NS31, previously used by Davison *et al.*, (2015), were used to amplify the 18S (SSU) rRNA gene of all isolates. This was done to observe whether this

marker could reveal genetic differences among the isolates used in this study and to compare those differences with the genetic variation revealed with ddRADseq data. The amplified PCR product was cloned with a StrataClone PCR Cloning Kit (Agilent Technologies ®) following manufacturer protocols and one to two clones were sequenced. For more information of PTG and SSU amplification, sequencing and sequence cleaning see Supplementary note S3.

ddRAD-sequencing

Double-digest restriction-site associated DNA sequencing was carried out on DNA of the 61 isolates with three to five independent biological replicates per isolate (Wyss *et al.*, 2016; Supplementary note S4). The sampling design is given in Supplementary note S5. The ddRAD-seq protocol was performed on 233 DNA samples. Between 1 and 25ng/µl of genomic DNA was digested with the *EcoRI* and *MseI* enzymes (NEB). Between 36 and 90 samples were pooled into each of four libraries with independent barcode for each sample. Library quality was verified on a Fragment Analyser® (Advanced Analytical Technologies). Each library was then sequenced at the Lausanne Genomics Technologies Facility (GTF) on an Illumina HiSeq sequencer in one lane to give 100bp paired-end reads.

In silico *workflow*

An additional ddRAD-seq dataset containing 20 isolates of *R. irregularis*, with 3 biological replicates of each isolate, was included in the analyses (Wyss *et al.*, 2016), thus increasing the total number of isolates analysed to 81 (and the number of samples analysed to 299). Read quality was controlled with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The computations were performed at the Vital-IT (http://www.vital-it.ch) center for high-performance computing of the Swiss Institute of Bioinformatics (SIB). After Illumina adapter and low quality read trimming and de-multiplexing (Supplementary note S6), the reads were mapped to the DAOM197198 *R. irregularis* genome assembly N6 with the Novoalign software (Novocraft-Technologies 2014) . We chose this assembly, because it was the most complete single nucleus genome assembly of this fungus to date (Lin *et al.*, 2014). FreeBayes v1.0.2 (Garrison and Marth, 2012) software was used to call single nucleotide polymorphisms (SNPs), insertions and deletions (indels) and multiple nucleotide polymorphisms (MNPs). The –p

option was set up to 10, assuming a potential number of alleles up to 10 and a minimum frequency of each allele of 10%. VCFfilter from the VCFLib library (Garrison and Marth 2012) was used to call positions with a phred quality >30. Markers falling in coding and non-coding regions were defined based on a prediction of coding regions with GeneMark-ES (Ter-hovhannisyan *et al.*, 2008; Wyss *et al.*, 2016). Positions identified as repeated elements in the genome, using RepeatModeler Open-1-0 (Smit and Hubley, 2008) and RepeatMasker Open-3.0 (Smit *et al.*, 1996), were removed from subsequent analyses (Wyss *et al.*, 2016). An *in-silico* double digestion *Eco*RI-*Mse*I in the original genome allowed identification of predicted fragments and all fragments that potentially mapped two or more times to the DAOM197198 genome. We did not consider samples with less than 450 000 reads after demultiplexing, and samples with less than 300 000 reads uniquely mapping to the N6 assembly for further analyses. This decreased the number of samples to 262 (Supplementary Table S2). The four libraries, RAD7, RAD8, RAD10B and RAD12 are available at NCBI (bioproject accession number PRJNA326895).

SNP datasets

Three datasets were generated and analyzed. All the positions included in these datasets were shared among all isolates, so that no missing data was present for any isolate or replicate. These positions were located in coding and non-coding regions and contained at least 10 reads of coverage. The first dataset included all 262 samples and included 498 mono-allelic and poly-allelic positions in total. This dataset allowed us to conduct analyses to study genetic variation among all isolates and all species included in this study. A second dataset included only one replicate of each isolate, where the replicate chosen to represent each isolate was the one with the highest number of uniquely mapped reads and the highest number of loci covered. This allowed us to conduct more precise analysis on the genetic differences among all isolates using a larger number of polymorphic sites than that available in dataset 1. Using this selection, a total of 2 491 mono-allelic positions were concatenated for 68 isolates. In this dataset, one isolate out of the 14 identified using the PTG sequences as *R. proliferus* was used as an outgroup. A third dataset comprised a single replicate of each isolate identified in the PTG tree as *R. irregularis*. By considering polymorphic positions that were only shared among all isolates identified as *R. irregularis*, a database with a larger number of polymorphic positions could be generated, allowing us to conduct analyses with a high resolution to identify and quantify genetic differences within the species *R. irregularis*. In this

dataset, 6 888 mono-allelic positions were concatenated. A summary of the three datasets is found in Table 1 and information about the samples used in each of the three datasets is found in Supplementary Table S2.

Table 1: The three datasets used for the phylogenetic analysis and population structure analysis

	No. of Samples	No. of Isolates	No. of Sites	Taxa included in dataset
Dataset 1	265	81	489	*R. intraradices, R. irregularis, Rhizophagus* sp. LPA8-CH3, *R. proliferus*
Dataset 2	68	68	2491	*R. intraradices, R. irregularis, Rhizophagus* sp. LPA8-CH3, *R. proliferus* (as outgroup)
Dataset 3	59	59	6888	*R. irregularis*

Phylogenetic analyses and 18S rRNA clustering

Phylogenetic analysis of the PTG is described in Supplementary note S3. After sequence cleaning, SSU cloned sequences of each isolate were clustered at 97, 98 and 99% similarity thresholds with the script *pick_otus* and the *uclust* clustering algorithm in the QIIME environment (Caporaso, J. G. et *al.*, 2010, Edgar 2010). Each clone was also blasted against the MaarjAM database (Öpik et *al.*, 2010) and the best VT hit was retained. The sequences used for the PTG phylogeny comprised 71 from this study and 25 from Sokolski *et al.* (2011). All PTG sequences used to reconstruct this phylogeny have been deposited in NCBI GenBank as accession numbers KY348541 - KY348610 and KY436236. The SSU sequences were deposited in the NCBI GenBank with accession numbers KY436237-KY436352.

Scalar distances calculated on the first ddRAD-seq dataset were analysed with the package ape, version 3.5 (Paradis *et al.*, 2003) in the R software 3.3.2. A dendrogram was built using the function *hclust* and *plot.phylo* (Wyss *et al.*, 2016). A bootstrapping method described by Wyss *et al.* (2016), was applied to calculate support values. The second and third datasets were analyzed with MrBayes (Huelsenbeck and Ronquist, 2001), implementing Markov chain Monte Carlo (MCMC) probability analysis. A mixed model was set up with the command lset nst=mixed rates=gamma. The MCMC chains were run for 1 000 000 generations with the reversible jump MCMC (RJ-MCMC) procedure, avoiding the selection of only one model of substitution rate. The chains were run until the standard deviation (SD) of split frequencies reached 0.01. All PSRF+ values reached a value of 1.

Population structure, species and clade delimitation

A constrained correspondence analysis (CCA) built with the function *cca* from the package vegan, version 2.4-1 (Dixon, 2003) was applied to the first dataset and based on a matrix of scalar distances among the 262 samples and the 498 positions. We did this to examine how all the different isolates used in this study would cluster.

BP&P (Yang and Rannala, 2010, version 3.3) and STRUCTURE (Pritchard *et al.*, 2000, version), two programs based on the Bayesian algorithm, were used in order to define species and genotype boundaries. First, the 2 491 positions of the second dataset were concatenated and tested under the multispecies coalescent (MSC) model (Rannala and Yang, 2003; Yang, 2002). Using this model, species were delimited using a user-specified guide tree (Yang and Rannala, 2010; Rannala and Yang, 2013). The BPP software only evaluates the models that can be generated by collapsing nodes on the guide tree using the reversible jumps algorithm. We have added this text into the manuscript. The specific guide tree used is based on the phylogeny generated in Mr Bayes, as follows:

(*R. proliferus*, (*R. intraradices*, (*Rhizophagus* sp. LPA8-CH3 , ((*R. irregularis* Gp3, *R. irregularis* Gp4), ((*R. irregularis* Gp1A, *R. irregularis* Gp1B), *R. irregularis* Gp2))))),

where *Rhizophagus* sp. LPA8-CH3 formed a distinct sister clade to *R. irregularis* comprising two isolates, *R. intraradices* is a known species, and *R. irregularis* groups Gp1A, Gp1B, Gp2, Gp3, Gp4 are distinct genetic groups within the phylogenetic analysis. Three models were run with varying ancestral population size (θ) and root age (τ_0) following the protocol of Leaché and Fujita (2010**)** (Supplementary note S7).

Population structure in the third dataset was analyzed using an admixture ancestry model with the program STRUCTURE, without prior knowledge of location and with correlated allele frequencies between populations. STRUCTURE (Pritchard *et al.*, 2000) is a program suitable for defining species groups (K) even when the Hardy-Weinberg equilibrium is not respected and is also robust for analysis of population structure with data originating from organisms with low recombination rates (Falush *et al.*, 2003, 2007). The MCMC chains were run for 100 000 generations, with 10 000 generations of burn-in. Each run from K1 to K7 was replicated 10 times. The delta K (Evanno *et al.*, 2005) was computed using STRUCTURE HARVESTER (Earl and vonHoldt, 2012) in order to define the most likely number of K present in the sample. In order to concatenate all the replicates for each K, CLUMPP was

used (Jakobsson and Rosenberg, 2007). Graphical output of cluster assignment as barplots and pie charts for mapping each isolate to their geographical origin was performed in R.

Finally, an analysis of molecular variance (AMOVA) was performed on the data using Continent as a factor with the highest number of isolates (i.e. Europe and North America), with 1000 permutations, based on scalar genetic distances (pegas package version 0.9, Paradis, 2010). A Mantel test with the package ape (Paradis *et al.*, 2004, version 3.5) and with 1000 permutation was also performed to test for isolation by distance, by correlating a geographical distance matrix and the matrix of phylogenetic scalar distance based on dataset 1 derived from ddRAD-seq.

A previous study (Croll *et al.*, 2008) analyzed the genetic structure among 29 *R. irregularis* isolates with a reduced set of markers comprising ten microsatellites, two mitochondrial LSU gene introns and one nuclear gene intron. Those 29 isolates were also present in this study, allowing a comparison of the results. We produced a matrix of phylogenetic scalar distance, based on the ddRADseq dataset, but this time with the 29 isolates. This matrix was then compared with a Mantel test to a matrix of phylogenetic Jaccard distance based on 13 markers (Croll et *al.*, 2008). This analysis combined with the other analyses indicated that the ddRADseq data were robust (Supplementary note S8).

Measurement of hyphal density

Previous studies have shown differences in hyphal density among some of the isolates used in this study. We wanted to see whether hyphal density patterns of *R. irregularis* varied in accordance with the genetic differences observed among *R. irregularis* isolates. After sequencing and analyzing all ddRADseq of each isolate, nine isolates spread across the major branches of the phylogeny were cultured for three months *in vitro* with five replicates. There were only three replicates of one of the isolates, DAOM197198-CZ. Following three months of growth, each plate was photographed using a camera attached to a stereomicroscope in order estimate hyphal density (Supplementary note S9).

Results

Phylogeny based on phosphate transporter gene sequences

Sequences of the PTG ranged between 800 to 1 191 bp in length, with 282 SNPs. Previously published sequences (Sokolski et *al.*, 2011) and sequences of the reference isolates FL208 and DAOM197198, combined with the other PTG sequences from this study, resolved three *Rhizophagus* species; *R. proliferus, R. intraradices* and *R. irregularis* (Figure 1a). Twenty-two different haplotypes were identified with a nucleotide diversity (PiT) of 0.0806. In the *R. irregularis* clade, 13 haplotypes were detected with a nucleotide diversity of 0.0158 and 49 segregating sites. As expected, sequences of the reference strain DAOM197198 clustered within the *R. irregularis* clade. A highly supported sister clade to *R. irregularis* was designated as a distinct putative species *Rhizophagus* sp. LPA8-CH3 based on the two isolates comprising this group. The isolate FL208, considered as the holotype of *R. intraradices* (Stockinger et *al.*, 2009), clustered with five other isolates and with a previously published sequence of *R. intraradices.*

Sequences of the same SSU region studied by Davison *et al.*, 2015 of 554 bp, revealed that at the 97% similarity threshold the different *Rhizophagus* species used in this study could not be discriminated from each other (Figure 1b; Supplementary Table S3). Only at the 99% similarity threshold, could the 3 species of *Rhizophagus* plus the putative species be distinguished from each other (Figure 1b). At the 99% similarity threshold, *R. proliferus*, *R. intraradices* and the putative *Rhizophagus* sp. LPA8-CH3, each occurred in one separate VT (as seen in the MaarjAM database). *R. irregularis* isolates occurred in three different VT.

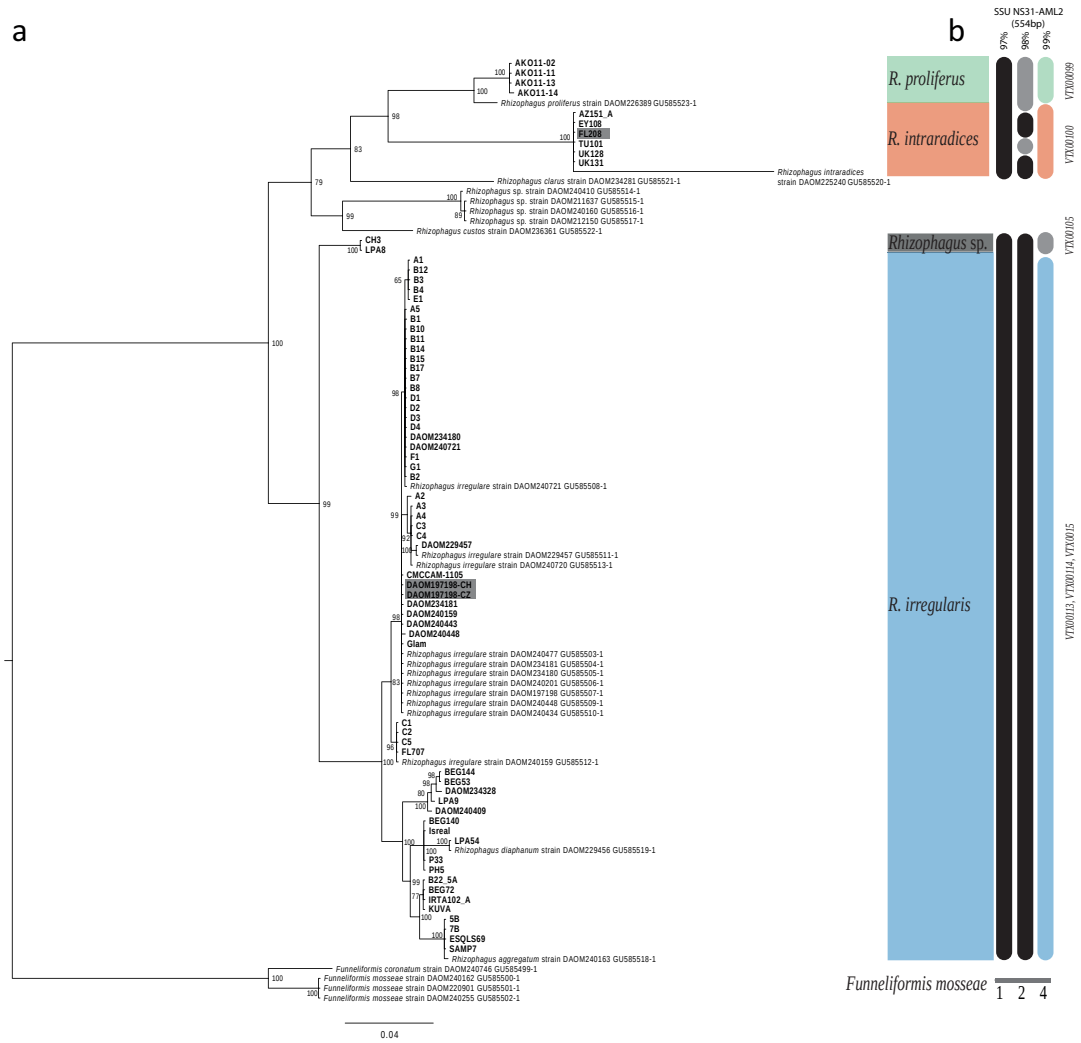


Figure 1. Phylogeny constructed using phosphate transporter gene (PTG) sequences and schematic clustering of the SSU. (a) Additional sequences from Sokolski *et al.*, 2011. Sequences of *Funneliformis mosseae* were used to root the tree. Holotypes of *R. intraradices* (FL208) and *R. irregularis* (DAOM197198) shown in dark grey. *R. proliferus* (green), *R. intraradices* (red), *R. irregularis* (blue) and *Rhizophagus* sp. LPA8-CH3 (dark grey). Numbers at nodes represent bootstrap support for consensus tree bipartitions. **(b)** SSU sequences for all isolates clustered at 97%, 98% and 99% and corresponding virtual taxa number (VT) of the first hit after blast on MaarjAM database.

Analysis of ddRAD-seq data

Each of the four ddRAD-seq libraries generated more than 300 million reads (Supplementary Table S4. Barcodes to de-multiplex the samples in each library are listed in Supplementary Table S2.

Sequencing details for each replicate of each isolate are reported in Supplementary Table S2. The ddRAD-sequencing on DNA from 81 isolates resulted in 509 563 positions covered with a mean of 12 282 SNPs and 1936 indels per isolate when compared to the reference genome

assembly. Not all of these positions had sufficient depth of coverage across all isolates in all samples. Among *R. intraradices* isolates, the number of positions in non-repeated coding and non-coding regions, covered by at least 10 reads, ranged from 23 668 in UK131-1 to 80 460 in UK128-1. Among *R. proliferus* isolates, the number of positions ranged from 60 795 in AKO11-11-2 to 77 186 in AKO11-13-1. Among *R. irregularis* isolates, positions ranged from 73 030 in ESQLS69-1 to 361 950 in DAOM240448-5.

Based on the phylogenies of both ddRADseq datasets 1 and 2, and a CCA built with the first dataset (Supplementary Figure S1, Figures 2 and 3), similar clade structure to that found with the PTG phylogeny was obtained: *R. proliferus* (n=14), *R. intraradices* (n=6) and *R. irregularis* (n=59). These data provided additional support for recognizing *Rhizophagus* sp. LPA8-CH3 (n=2) as being genetically distinct from *R. irregularis* and *R. intraradices*. Four mains branches within the *R. irregularis* tree as well as four main groups of isolates within *R. irregularis* were clearly recognized in the CCA.

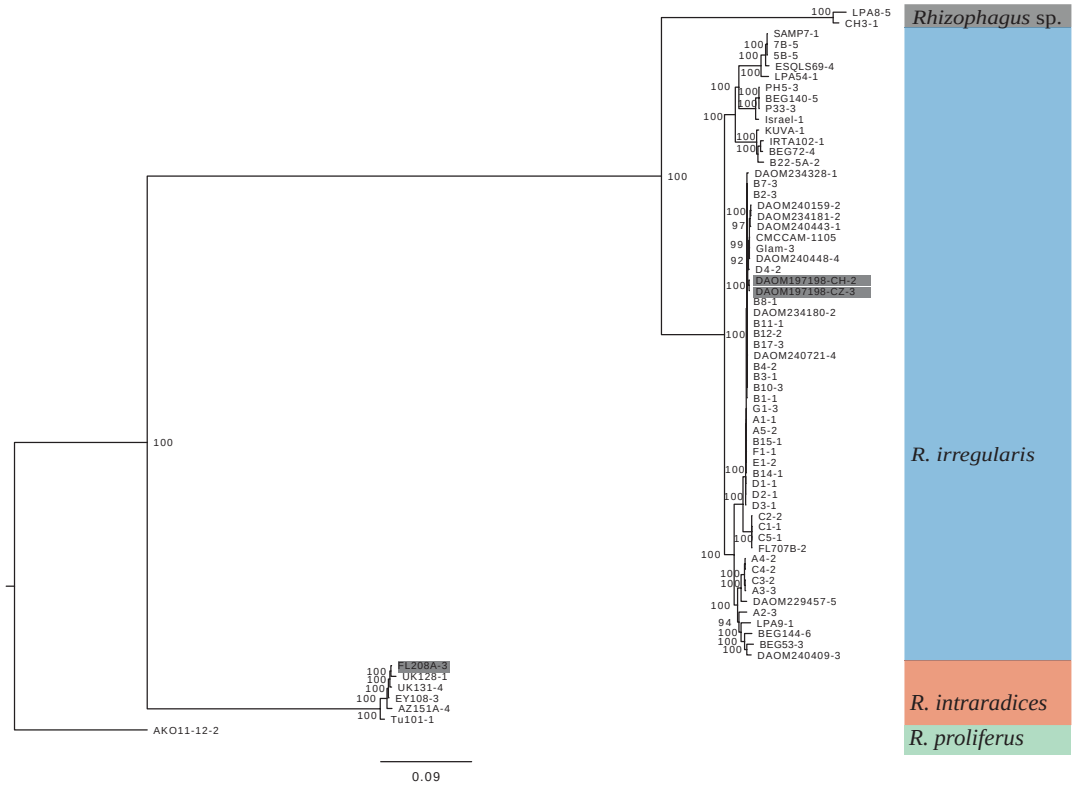


Figure 2. Phylogenetic tree based on a concatenation of 2 491 SNPs across the genome (dataset 2) of three species of *Rhizophagus* and *Rhizophagus* sp. LPA8-CH3 and 68 isolates. Only one replicate per isolate was used. Names are composed of the isolate name followed by the replicate number (short name in Table S2). Numbers at nodes represent bootstrap support for consensus tree bipartitions. Colour coding follows that of Figure 1.

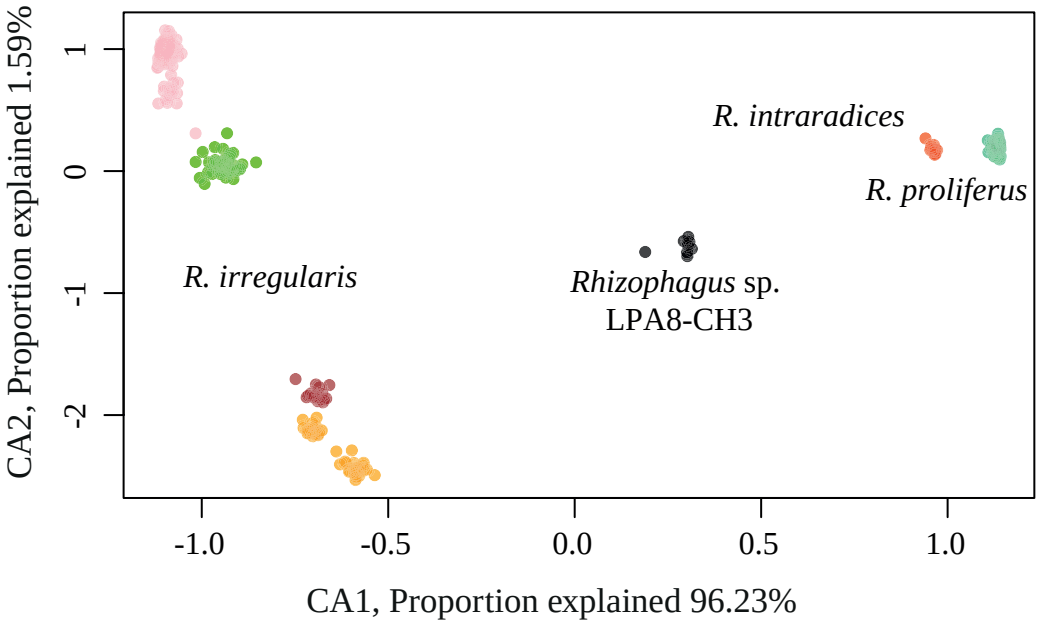


Figure 3: Constrained correspondence analysis (CCA) built on a scalar distance matrix computed among all samples and replicates (n=262) and based on 489 shared mono-allelic or poly-allelic positions. The four AMF species are well separated along the X-axis. Colour coding follows that of previous figures.

Phylogeography and population genetic structure

Rhizophagus irregularis and *R. intraradices* isolates co-occurred in Europe, the Middle East, North Africa and Northern America (Supplementary Figure S2.). We analyzed the speciation probability among *Rhizophagus* species as well as within *R. irregularis*. The MSC model implemented with dataset 2 using BP&P showed that in three models, the three distinct species were highly supported: *R. irregularis, R. intraradices* and *R. proliferus* and one putative species: *Rhizophagus* sp. LPA8-CH3 (Figure 4). A probability of > 0.95 would normally be considered likely to represent a speciation event. In this analysis we also found significant divergence between groups Gp3 and Gp4 and between groups Gp1 and Gp2 within *R. irregularis*. Two of the three tests exceeded the > 0.95 probability threshold for genetic differentiation between Gp1A and Gp1B.

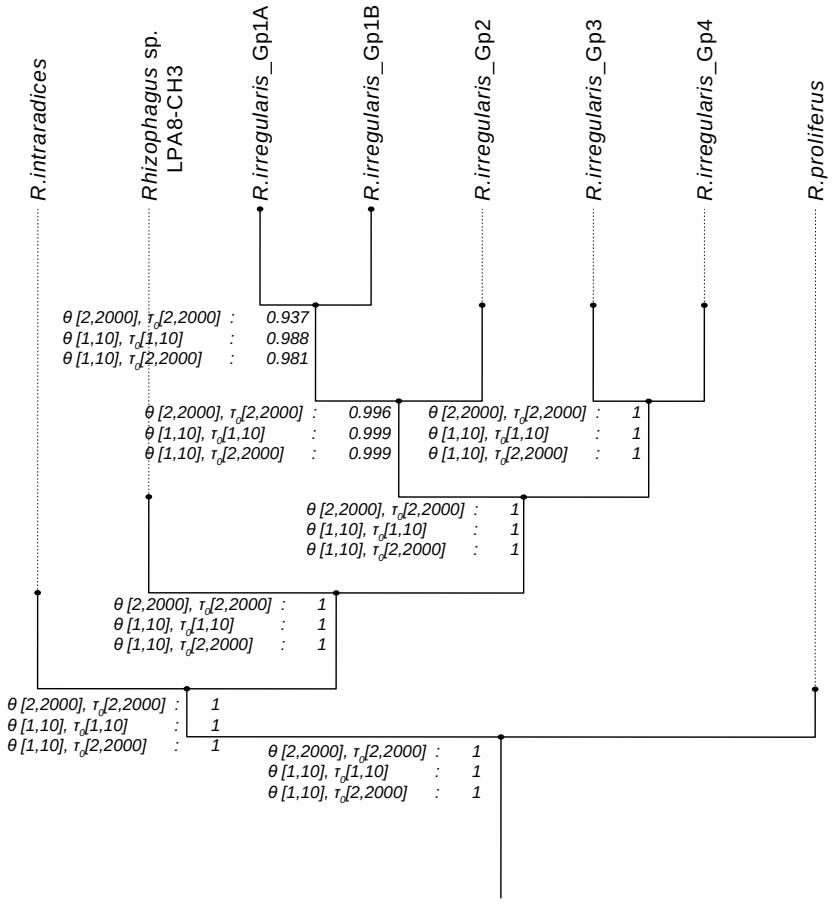


Figure 4: Bayesian species delimitation tree (based on dataset 2). Marginal probabilities of speciation for three models with variable population size (θ) and divergence time (τ0) are presented at each node.

The analysis using STRUCTURE and the Delta K on dataset 3 (59 *R. irregularis* isolates) suggested the possibility of either two or four divergent clades within the sampled *R. irregularis* isolates (Figure 5a), with Delta K values supporting a slightly higher probability of four groups than two (Supplementary Figure S3). BP&P and STRUCTURE analyses collectively suggested that the *R. irregularis* isolates were divided into at least four well-defined genetic groups.

The distribution of the divergent genetic groups of *R. irregularis* did not follow a geographical pattern (Figure 5b). Indeed, the Mantel test used to test for isolation by distance and an AMOVA using genetic distance and Continents as factor (America=9, Europe=37) did not reveal significant genetic differences (AMOVA: df=1, p=0.395) between the populations of the two different continents or any significant isolation by distance (z-statistic: 112835, p-value: 0.518). Moreover, isolates of each of the four genetic groups could be found in highly distant locations. For example, *R. irregularis* group Gp4 occurred in several geographically

distant places in north America (Florida and Canada), in Europe from Switzerland to Finland and in North Africa (Tunisia). Similar patterns were evident in the other genetic groups. Two strains (DAOM234181 and DAOM240159) appeared to be the same clonal haplotype and originated from two locations separated by almost 4000 km in Canada. Isolates belonging to the two close genetic groups Gp3 and Gp4 coexisted in the same soil in Switzerland and both groups also co-occurred in geographically close sampling sites in Canada.

The population structure of 29 *R. irregularis* isolates measured using 13 markers (Croll *et al.*, 2008) reflected that observed with the large number of SNPs identified in this study. The Jaccard distances among isolates measured with microsatellite data were significantly correlated with the scalar distance calculated on 2100 positions from ddRAD-seq data (Supplementary Figure S4, Mantel statistic r: 0.9624, p-value < 0.001).

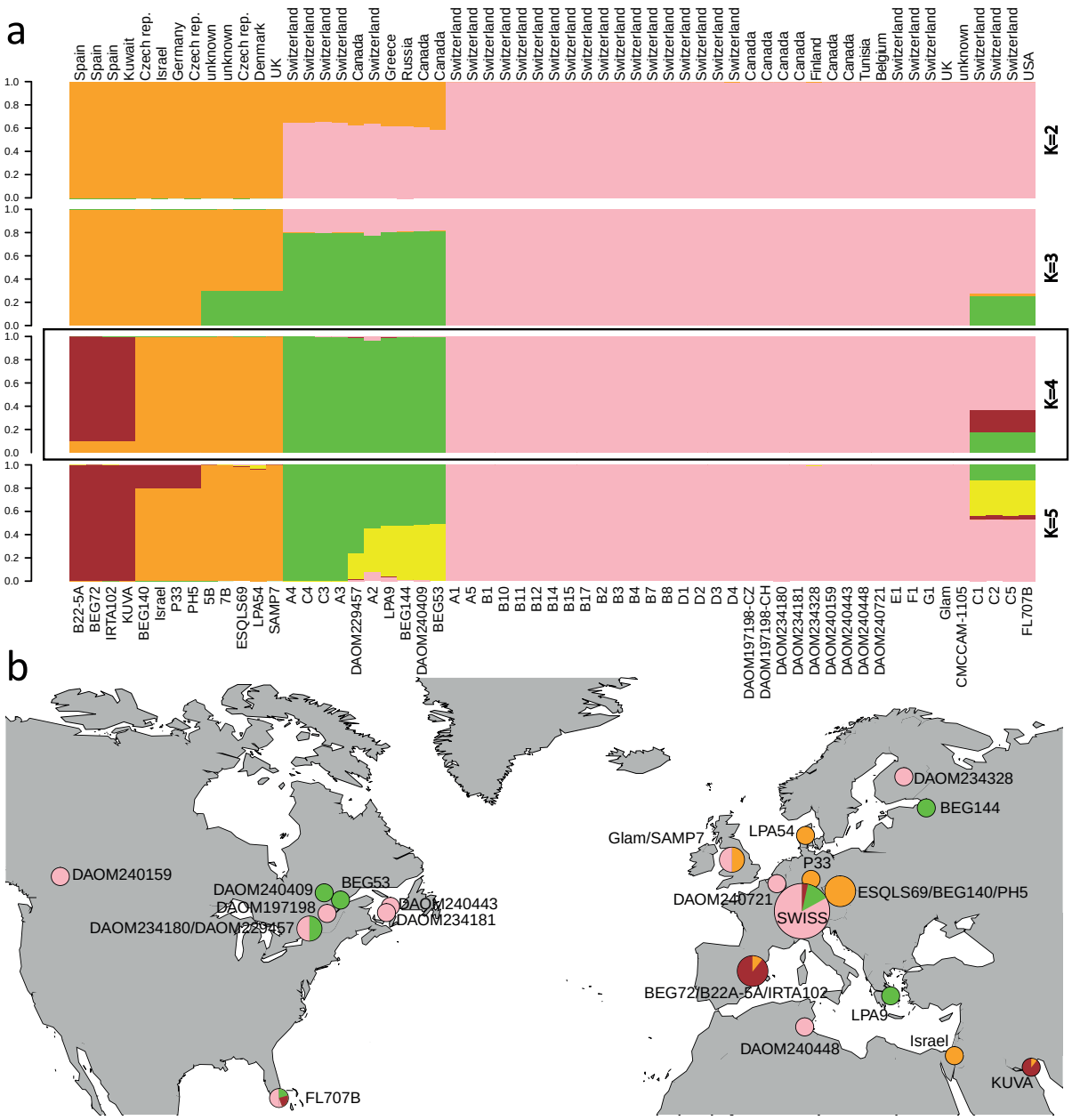


Figure 5: Assignment to 4 genetic groups with STRUCTURE and distribution map of genetic groups. (a) Vertical bars represent one *R. irregularis* isolate of and colours represent the assignment to different genetic groups (K). The delta K value for K = 4 was slightly higher than for K = 2. **(b)** Colour coding corresponds to results obtained with STRUCTURE, based on dataset 3. Several isolates are of unknown geographical origin and are not presented here. The name of each isolate is written near its respective pie chart, except for the 29 Swiss isolates that are represented as the Swiss population.

Extra-radical hyphal density

Dataset 3 provided the finest resolution to discriminate between *R. irregularis* isolates and showed, based on 6 888 SNPs, that genetic variation within each of the 4 main genetic groups also existed (Figure 6a). Significant differences in extraradical hyphal density among the nine

chosen *R. irregularis* isolates, and among three genetic groups (Gp1, Gp3 and Gp4), was found (Figure 6b; lmer; Genetic groups: dF = 2, F.value = 5.94, p = 0.035*). Isolates in the genetic group Gp3 produced a significantly higher density of extraradical hyphae than those in Gp4. Other comparisons between genetic groups were non-significant (Gp 3 - 4, p= 0.01*, Gp 1 - 4, p= 0.09 ., Gp 1 - 3, p= 0.18).

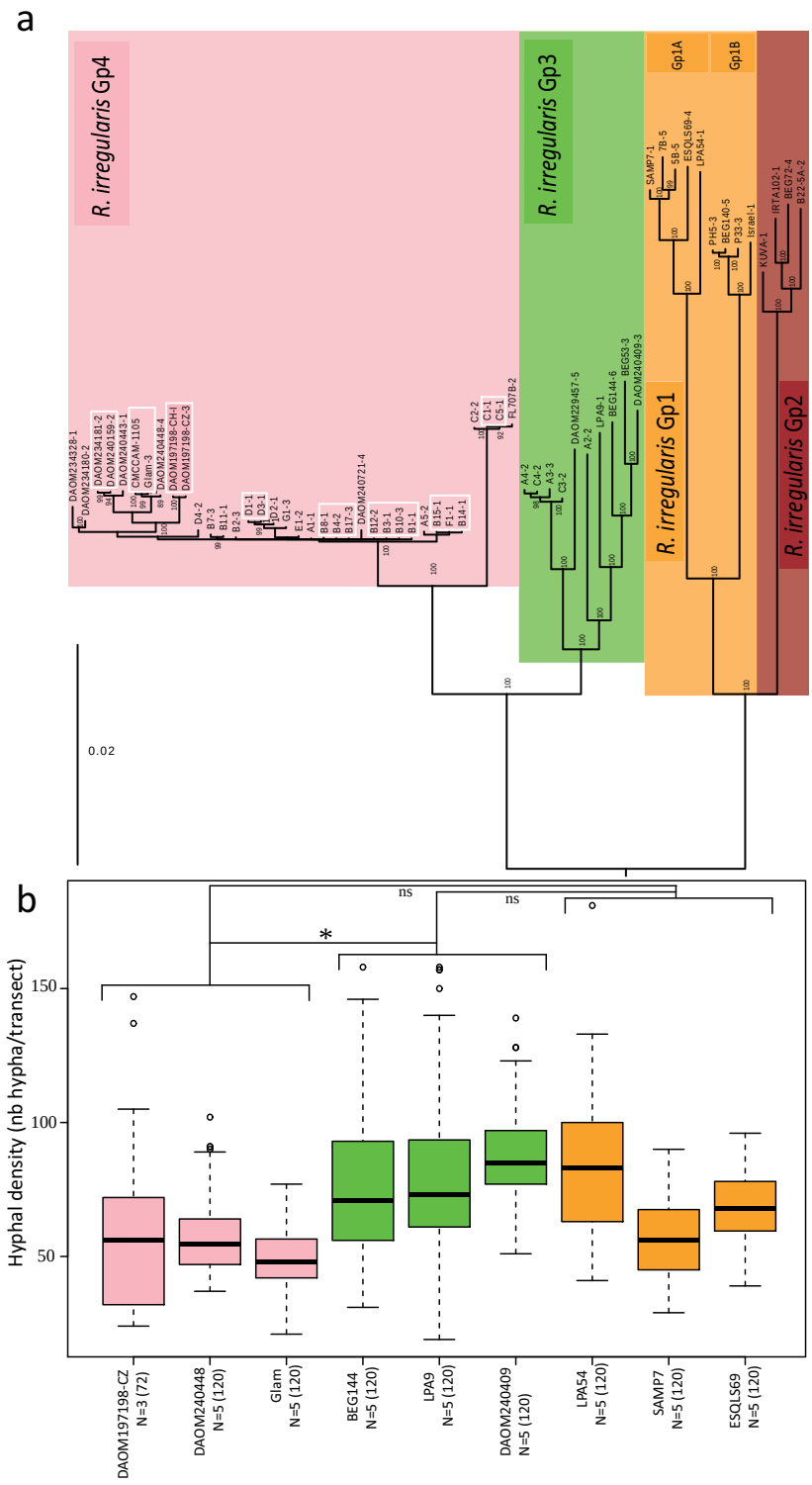


Figure 6: Phylogeny of the four genetic groups and hyphal density of 9 *R. irregularis* isolates from three genetic groups. Colour coding follows that in previous figures and corresponds to the four genetic groups. **(a)** Phylogeny based on 6 888 concatenated SNPs across the *R. irregularis* genome constructed using data from 59 isolates (dataset 3). The white squares represent clonal haplotypes. **(b)** Extra-radical hyphal density of *R. irregularis* isolates from three genetic groups. Genetic group 2 (Gp2) was not included. On the x-axis, N represents the number of plates and the number in parentheses represents the number of transects where the hyphae were counted. The significance among the three genetic groups and obtained with the mixed model is indicated above the boxplots (ns: non significant, * : p-value<0.05).

Discussion

Describing the genetic diversity of AMF species, and especially that of the broadly distributed *R. irregularis* group, is fundamental to understanding their biogeographic patterns. Data confirmed that both *R. irregularis* and *R. intraradices* are widely distributed geographically across at least two continents (North America and Europe). Moreover, the four cryptic genomic forms within the species *R. irregularis*, identified by thousands of SNP markers obtained in this study, also have a broad geographic distribution. These data collectively support the hypothesis of Davison *et al.*, (2015) that endemism is low in some Glomeromycota, not only at the 97% level of clustering resolution using VT and at the species level, but also amongst significantly diverged genomic forms within a species.

We confirmed with both the PTG and ddRAD-seq loci that the two reference isolates FL208 and DAOM197198 (holotypes of *R. intraradices* and *R. irregularis*, respectively) indeed belonged to two distinct clades and that a number of the other isolates in this study also clustered into these two clades. The four distinct genetic groups identified by variation in thousands of genome-wide SNPs among *R. irregularis* isolates were not fully resolved by the PTG and SSU phylogenies. This result stresses the need for higher resolution and larger number of markers for understanding AMF biogeography.

While the four genetically defined *R. irregularis* groups cannot be fully ranked at the species level in this study, the data indicate that, at most, a negligible amount of gene flow occurs among the groups, even though they coexist in the same soil. Thus, rather than referring to these as different cryptic species, we define the genetically different groups as cryptic genomics forms of the fungus, meaning that they are genetically distinct, but would likely not be distinguishable by morphological studies or using single gene markers. Anastomosis has been observed in the laboratory at very low frequency between isolates in groups Gp3 and Gp4 originating from the same location (Croll et *al.*, 2009). However, if this were common in nature, then such defined genetic groups would not be expected in the datasets generated in this study because intermediates should occur.

The four genetic groups were found across large geographical distances. The most striking examples were genetic groups Gp3 and Gp4, which occur at several different localities in Europe, Northern Africa, Canada and the USA. Some isolates from the same genetic group occurred more than 8000 km apart. The four genetic groups described here were detected

because of the high resolution obtained by the large number of ddRAD-seq markers. Thus, Bruns and Taylor (2016) were correct that low-resolution markers such as the one used by Davison *et al.*, (2015) would have failed to detect such genetic groups and would have assigned them to only one VT. However, regardless of the level of resolution informed by the SNPs or the 18S (SSU) rRNA, the same pattern of low endemism observed by Davison *et al.*, (2015) was evident for *R. irregularis*. This suggests, as proposed by another study (Rosendahl *et al*, 2009), that at least the AMF species studied in detail so far, from the species level down to the intraspecific level, have a wide distribution attributable either to high dispersal via anthropogenic activities (Davison et *al.*, 2015) or by slower dispersal over millions of years (Morton, 1990). Since many of the fungi in this study were isolated from agricultural fields, human agriculture may well be responsible for the dispersal of this species. We do not, however, consider it likely that the wide distribution of some of the groups is due to the application of exotic inoculum containing *Rhizophagus irregularis* because the more widespread use of such inocula is actually very recent and many of the donated cultures were isolated from the soil long before the use of commercial inocula.

The considerable genomic variability described here for a large set of isolates has not previously been reported for *R. irregularis*. Given that genetically different AMF isolates can cause large differences in plant growth (Munkvold *et al.*, 2004, Koch *et al.*, 2006), it would now be important to identify if variation in plant growth during the symbiosis with these different forms is greater among or within *R. irregularis* genetic groups. In this present study, we found evidence that the extraradical hyphal density, a phenotypic trait that is known to impact the phosphate acquisition capacity of the fungus and benefit to the plant (Jakobsen *et al.* 1992), differed significantly among genetic groups. Identifying whether the within species genetic and phenotypic variability in these fungi has consequences for plant ecology would be an important step to understand the link between fungal communities and plant communities as well as an important step for using AMF in agriculture for increasing crop yields.

The focus on the biogeography of *R. irregularis* in this study has broad agronomic implications. Genetically different *R. irregularis* applied to rice (Angelard *et al.*, 2010) resulted in differential rice growth. Significant yield increases in cassava in the field have been achieved by applying *in vitro* produced *R. irregularis* (Ceballos *et al.*, 2013). *Rhizophagus irregularis* is a potentially strong candidate species for large-scale inoculation of tropical crops because it can be mass-produced in contaminant-free *in vitro* conditions and

because of its global distribution. The occurrence of *R .irregularis* genotypes distributed across diverse environments suggests broad adaptability to different soil types.

The introduction of non-native exotic strains with a strong geographically determined genetic structure among populations could be perceived as a risk for potential invasiveness (Rodriguez and Sanders, 2015; Schlaeppi *et al.*, 2016). However, that concern is alleviated for *R. irregularis*, given that distribution of genetically similar isolates is widespread. For example, if isolates Gp3 and Gp4 from Switzerland were introduced in a field sites in Canada, the presence of these genotypes there negate any classification as exotic. We propose that a population genomic study of *R. irregularis* from tropical soils should be undertaken to verify if similar patterns occur.

Attention then would focus on the design of multi-isolate commercial inocula, where relatedness of isolates growing on the same plant strongly impacts plant biomass (Roger *et al.*, 2013). The variation found in the 81 isolates of this study provides a bank of genetic diversity that can be utilized to test for crop compatibility and symbiotic efficiency and effectiveness. The variability described in this study has the advantage that each isolate is referenced and stored as *in vitro* pure culture and could be at any time mass-produced and used for breeding programs (Rodriguez and Sanders, 2015), ecological experiments (Koch *et al.*, 2006) and agronomic trials (Ceballos *et al.*, 2013).

This study is, to our knowledge, the first to show that almost-clonal isolates of *Rhizophagus irregularis* occurred in highly distant localities up to 4000 km apart. It also confirms the findings of Davison et *al.*, (2015) of low endemism and similar genomic forms on multiple continents. At the same time, our results also show that, as Bruns and Taylor (2016) suggested, considerable genetic divergence can indeed be hidden by the use of low-resolution markers. The presence of different genetic groups of *R. irregularis* across the globe should be now investigated by isolating and genotyping isolates from agronomic and natural ecosystems from other continents such as Australia, Africa, Asia and South America.

Acknowledgements

We thank Aleš Látr and Miroslav Vosatka at Symbiom, Dirk Redecker, Amelia Campubri, Yolande Dalpé, Alok Adholeya and Louis Mercy for providing AMF isolates. We thank Jeremy Bonvin, Marie Strehler, Laura Alvarez Lorenzana, Jérôme Wassef, Jérôme Goudet, Lucas Villard and Nicolas Salamin for their help and three anonymous reviewers for their constructive comments. We thank Keith Harshman at the Lausanne University Genomic Technologies Facility. All bioinformatics computations were performed at the Vital-IT (http://www.vital-it.ch) Center for High Performance Computing of the Swiss Institute of Bioinformatics. This study was funded by the Swiss National Science Foundation (Grant number: 31003A_162549 to IRS).

References

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:** 1216-1221.

Baas Becking, L.G.M. (1934) Geobiologie of inleiding tot de milieukunde. The Hague, the Netherlands: W.P. Van Stockum & Zoon (in Dutch).

Bruns TD, Taylor JW (2016). Comment on "Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism". *Science* **351**.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7:** 335-336.

Ceballos I, Ruiz M, Fernandez C, Pena R, Rodriguez A, Sanders IR (2013). The *in vitro* mass-produced model mycorrhizal fungus, *Rhizophagus irregularis*, significantly increases yields of the globally important food security crop cassava. *Plos One* **8**.

Croll D, Wille L, Gamper HA, Mathimaran N, Lammers PJ, Corradi N *et al* (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:** 672-687.

Croll D, Giovannetti M, Koch AM, Sbrana C, Ehinger M, Lammers PJ et al. (2009). Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus *Glomus intraradices*. *New Phytol* **181**: 924–937.

Davison J, Moora M, Ouml, pik M, Adholeya A, Ainsaar L *et al* (2015). Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science* **349:** 970-973.

Dixon P (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14:** 927-930.

Earl DA, Vonholdt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4:** 359-361.

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

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611-2620.

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164:** 1567-1587.

Falush D, Stephens M, Pritchard JK (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7:** 574-578.

Finlay BJ (2002). Global dispersal of free-living microbial eukaryote species. *Science* **296:** 1061-1063.

Garrison E, Marth G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv:1207.3907 [q-bio.GN].

Huelsenbeck JP, Ronquist F (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754-755.

Jakobsen I, Abbott LK, Robson AD. (1992). External hyphae of vesicular-arbuscular mycorrhizal fungi associated with *Trifolium subterraneum* L. 1. Spread of hyphae and phosphorus flow into roots. *New Phytologist* **120:** 371-380.

Jakobsson M, Rosenberg NA (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23:** 1801-1806.

Koch AM, Croll D, Sanders IR (2006). Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. *Ecology Letters* **9:** 103-110.

Leaché AD, Fujita MK (2010). Bayesian species delimitation in West African forest geckos (Hemidactylus fasciatus). *Proceedings of the Royal Society B-Biological Sciences* **277:** 3071-3077.

Lin K, Limpens E, Zhang ZH, Ivanov S, Saunders DGO, Mu DS *et al* (2014). Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. *Plos Genetics* **10**.

Margaritopoulos JT, Kasprowicz L, Malloch GL, Fenton B. (2009). Tracking the global dispersal of a cosmopolitan insect pest, the peach potato aphid. *BMC Ecol* **9**: 13.

Mensah JA, Koch AM, Antunes PM, Kiers ET, Hart M, Bucking H (2015). High functional diversity within species of arbuscular mycorrhizal fungi is associated with differences in phosphate and nitrogen uptake and fungal phosphate metabolism. *Mycorrhiza* **25:** 533-546.

Monti F, Duriez O, Arnal V, Dominici JM, Sforzi A, Fusani L *et al* (2015). Being cosmopolitan: evolutionary history and phylogeography of a specialized raptor, the Osprey *Pandion haliaetus*. *BMC Evolutionary Biology* **15**.

Morton JB (1990). Species and clones of arbuscular mycorrhizal fungi (Glomales, zygomycetes) - their role in macroevolutionary and microevolutionary processes. *Mycotaxon* **37:** 493-515.

Munkvold L, Kjoller R, Vestberg M, Rosendahl S, Jakobsen I (2004). High functional diversity within species of arbuscular mycorrhizal fungi. *New Phytologist* **164:** 357-364.

Novocraft-Technologies. (2014). Novoalign. Available from http://www.novocraft.com

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

Opik M, Davison J, Moora M, Partel M, Zobel M (2016). Response to Comment on "Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism". *Science* **351**.

Paradis E, Claude J, Strimmer K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

Paradis E (2010). pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* **26:** 419-420.

Parchman TL, Gompert Z, Mudge J, Schilkey FD, Benkman CW, Buerkle CA. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Mol Ecol* **21**: 2991–3005.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012). Double Digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *Plos One* **7**: e37135.

Pringle A, Baker DM, Platt JL, Wares JP, Latge JP, Taylor JW (2005). Cryptic speciation in the cosmopolitan and clonal human pathogenic fungus *Aspergillus fumigatus*. *Evolution* **59:** 1886-1899.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155:** 945-959.

Rannala B, Yang ZH (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* **164:** 1645-1656.

Rannala B, Yang ZH (2013). Improved reversible jump algorithms for bayesian species delimitation. *Genetics* **194:** 245-253.

Rodriguez A, Sanders IR (2015). The role of community and population ecology in applying mycorrhizal fungi for improved food security. *ISME Journal* **9:** 1053-1061.

Roger A, Colard A, Angelard C, Sanders IR (2013). Relatedness among arbuscular mycorrhizal fungi drives plant growth and intraspecific fungal coexistence. *ISME Journal* **7:** 2137-2146.

Rosendahl S, McGee P, Morton JB (2009). Lack of global population genetic differentiation in the arbuscular mycorrhizal fungus *Glomus mosseae* suggests a recent range expansion which may have coincided with the spread of agriculture. *Molecular Ecology* **18:** 4316-4329.

Sanders IR (2010). 'Designer' mycorrhizas?: Using natural genetic variation in AM fungi to increase plant growth. *ISME Journal* **4:** 1081-1083.

Schlaeppi K, Bender SF, Mascher F, Russo G, Patrignani A, Camenzind T *et al* (2016). High-resolution community profiling of arbuscular mycorrhizal fungi. *New Phytologist* **212:** 780-791.

Smit AFA, Hubley R. (2008). RepeatModeler Open-1.0. http://www.repeatmasker.org.

Smit AFA, Hubley R, Green P. (1996). RepeatMasker Open-3.0. http://www.repeatmasker.org.

Smith SE, Read DJ. (2008). Mycorrhizal Symbiosis 3rd ed. Elsevier Ltd.: London, UK.

Sokolski S, Dalpe Y, Piche Y (2011). Phosphate transporter genes as reliable gene markers for the identification and discrimination of arbuscular mycorrhizal fungi in the genus *Glomus*. *Applied and Environmental Microbiology* **77:** 1888-1891.

St-Arnaud M, Hamel C, Vimard B, Caron M, Fortin JA (1996). Enhanced hyphal growth and spore production of the arbuscular mycorrhizal fungus *Glomus intraradices* in an *in vitro* system in the absence of host roots. *Mycological Research* **100:** 328-332.

Stockinger H, Walker C, Schussler A (2009). '*Glomus intraradices* DAOM197198', a model fungus in arbuscular mycorrhiza research, is not *Glomus intraradices*. *New Phytologist* **183:** 1176-1187.

Talas F, McDonald BA (2015). Genome-wide analysis of *Fusarium graminearum* field populations reveals hotspots of recombination. *BMC Genomics* **16**.

Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, Hibbett DS *et al* (2000). Phylogenetic species recognition and species concepts in fungi. *Fungal Genetics and Biology* **31:** 21-32.

Taylor JW, Turner E, Townsend JP, Dettman JR, Jacobson D (2006). Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. *Philosophical Transactions of the Royal Society B-Biological Sciences* **361:** 1947-1963.

Tedersoo L, Bahram M, Polme S, Koljalg U, Yorou NS, Wijesundera R *et al* (2014). Global diversity and geography of soil fungi. *Science* **346:** 1078-+.

Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M (2008). Gene prediction in novel fungal genomes using an *ab initio* algorithm with unsupervised training. *Genome Research* **18:** 1979-1990.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T *et al* (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396:** 69-72.

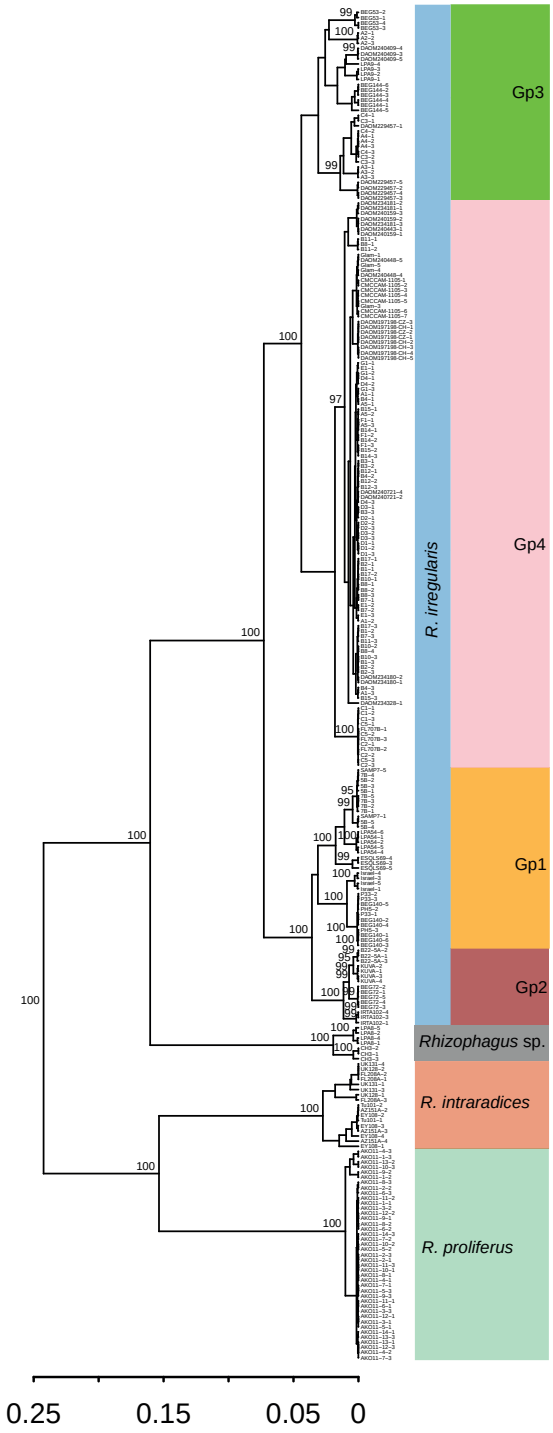
van der Heijden MGA, Martin FM, Selosse MA, Sanders IR (2015). Mycorrhizal ecology and evolution: the past, the present, and the future. *New Phytologist* **205:** 1406-1423.

Wyss T, Masclaux FG, Rosikiewicz P, Pagni M, Sanders IR (2016). Population genomics reveals that within-fungus polymorphism is common and maintained in populations of the mycorrhizal fungus *Rhizophagus irregularis*. *ISME Journal* **10:** 2514-2526.

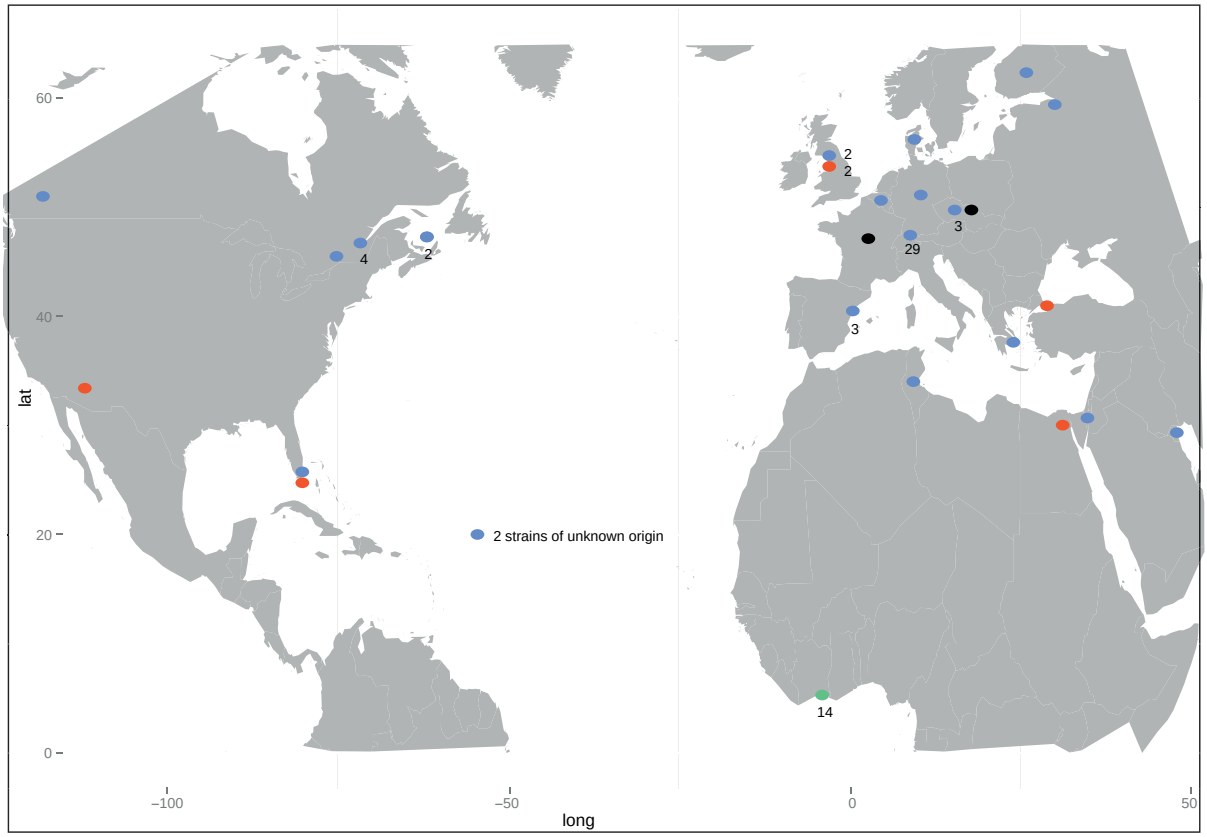
Yang ZH (2002). Likelihood and Bayes estimation of ancestral population sizes in hominoids using data from multiple loci. *Genetics* **162:** 1811-1823.

Yang ZH, Rannala B (2010). Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci USA* **107:** 9264-9269.

Supplementary information

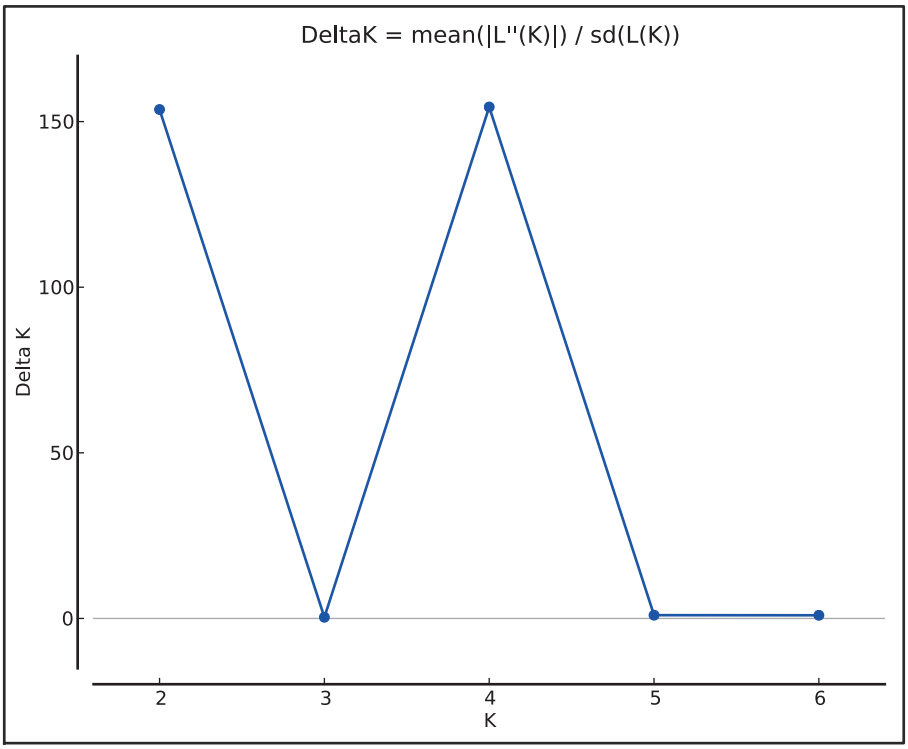


Supplementary Figure S1: Scalar distance tree based on 489 mono-allelic and poly-allelic markers of 81 isolates and their biological replicates. Three different species are depicted in blue (*R. irregularis*), in pale red (*R. intraradices*) and in pale green (*R. proliferus*). Additionally, *Rhizophagus* sp. LPA8-CH3 is depicted in grey. Within *R. irregularis*, different colours (pink: Gp4, green: Gp3, red: Gp2, orange: Gp1) correspond to four different genetic groups.

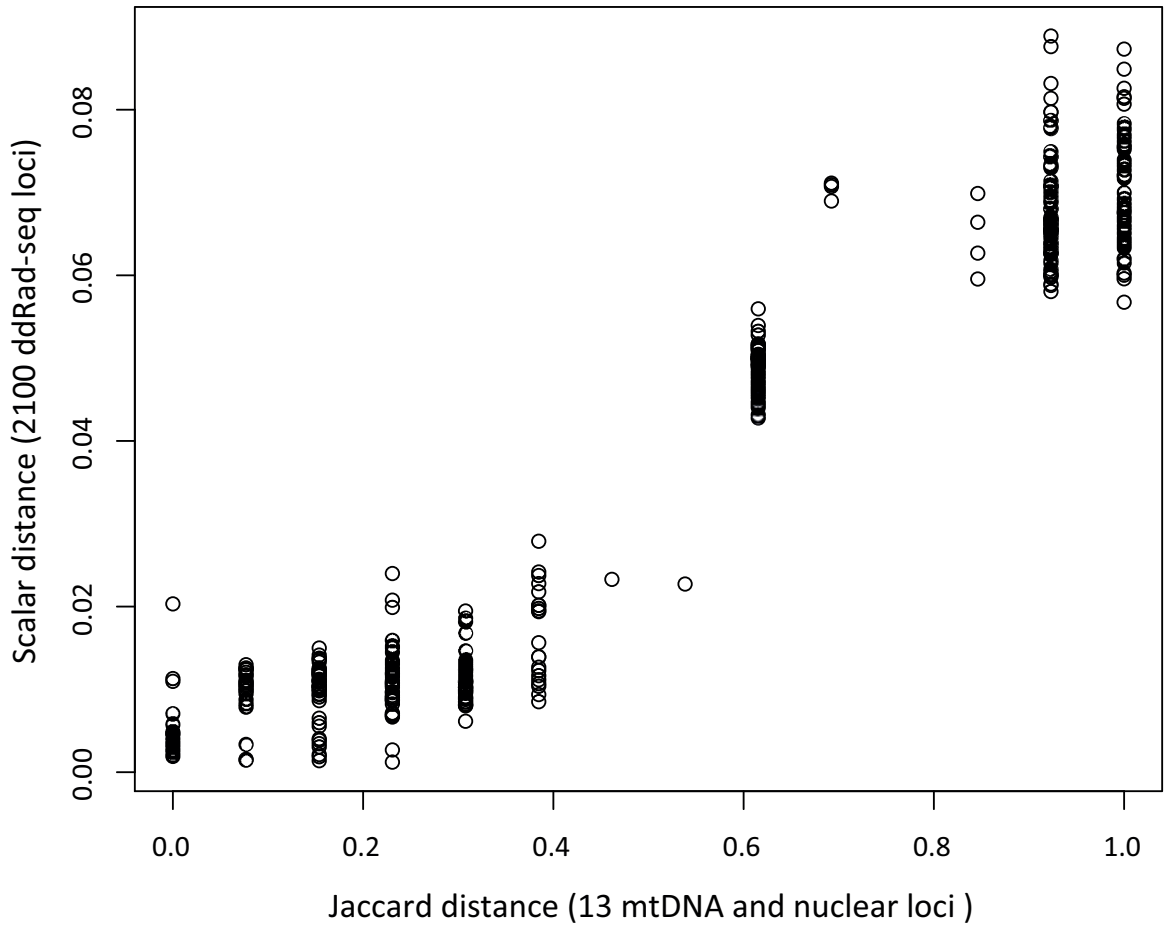


Supplementary Figure S2: Map of AMF species distribution

Geographical origin of isolates and species of AMF. In blue, *R. irregularis*, in red *R. intraradices*, in green *R. proliferus* and in black *Rhizophagus* sp. LPA8-CH3. Numbers correspond to the number of isolates isolated in each site. When no number is present, only one isolate originated from the corresponding location.



Supplementary Figure S3: StructureHARVESTER plot. The delta K (Evanno et al. (2005) suggests the best number of K. The delta K for K2 and K4 have the highest values with a slightly higher value for K4.



Supplementary Figure S4: Mantel correlation between Jaccard distance based on microsatellites, and scalar distance based, on ddRAD-seq data. Twenty-nine isolates that were previously genotyped by Croll *et al.* in 2008 were also analysed by ddRAD-seq in this study. The genetic distances between the isolates based on microsatellites, and calculated with Jaccard distance, were strongly correlated with the distance between isolates based on ddRAD-seq data, based on a scalar distance (Mantel statistic r: 0.9624, p-value<0.001)

Supplementary note S1: *Rhizophagus irregularis* nomenclature

Previously this species was ascribed as *Glomus intraradices* or was given the name *Glomus irregulare* (Stockinger *et al.,* 2009; Krüger *et al.,* 2011) or *Rhizoglomus irregulare* (Sieverding et *al.,* 2014). It was recently defined as a separated species to *Rhizophagus intraradices* mainly based on the two holotypes, respectively DAOM197198 and FL208 (Stockinger *et al.,* 2009), both of which are included in this study.

Supplementary note S2: Single spore isolation, decontamination and *in vitro* culturing protocol

Twenty g of soil containing AMF spores was sieved with 3 mesh sizes: 200, 50 and 32µm. The contents of the 32 and 50 µm sieves were centrifuged for 4 min in a 50ml falcon tube with water. A second step of centrifugation for 4 min was carried out, replacing 50% of the water with a 50% sucrose solution. Around 200 spores were then collected from the supernatant and decontaminated in Chloramid T for 4 min and rinsed 3 times with a gentamycin and streptomycin solution at 1% in a sterile hood. Clean spores were then placed on M medium, separated from each other on a grid. Successfully germinated single spores were transferred to a medium with transformed carrot roots. Each fungus was then grown in the *in vitro* culture system for 3 months in the presence of root tumor-inducing plasmid T-DNA-transformed carrot roots (Bécard and Fortin 1988).

Supplementary note S3: Phosphate transporter gene (PTG) and 18S (SSU) rRNA gene sequencing and analysis

The Phosphate Transporter Gene (PTG) was considered a phylogenetically informative marker by Sokolski et *al.,* 2011. We amplified this marker from DNA of each isolate using primers P4 and P6 (Sokolski *et al.,* 2011). We then sequenced the amplified product to confirm the species identity of each isolate. A PCR mix of 10µl was prepared with 1X of QIAGEN PCR Buffer, 1mM of $MgCl_2$, 0.2 mM of dNTP mix containing all dNTPs, 0.5 µM of each primer, 0.4 U of *Taq* polymerase (QIAGEN) and 2µl of gDNA at a concentration ranging between 1 and 25ng/µl. PCR conditions for amplification of the PTG on a Biometra T1 Thermocycler PCR machine were as follows: 5 min at 95°C followed by 35 cycles at 95°C for 10s, 50°C for 30s and 72°C for 2min, a final elongation step at 72°C for 10min.

The 18S (SSU) rRNA gene is a commonly used maker for NGS analysis of soil fungal communities; we used the primers AML2 and NS31 (Davison et *al.,* 2015) to amplify this

region. The PCR mix of 12.5µl for the amplification of the SSU region consisted of 1.56X of QIAGEN PCR Buffer, 0.0625 mM of dNTP mix containing all dNTPs, 0.3125 µM of each primer, 1.25 U of *Taq* polymerase (QIAGEN) and 1µl of DNA of the same extraction used for the PTG amplification. Each successfully amplified fragment of each isolate was then cloned with the StrataClone PCR Cloning Kit (Agilent Technologies ®) following manufacturer protocols. Two clones of each isolate were selected and amplified with the set of primers T3/T7.

Purification and Sanger sequencing of both markers was performed by GATC biotech (Germany). Quality assessment of the sequences, alignment and sequence cleaning was performed using MEGA 6.06 (Tamura *et al.,* 2013). MrBayes software (Huelsenbeck *et al.,* 2001) was used to infer genealogy of the PTG. The GTR+G substitution model was selected using the best AIC values from jModelTest 2.1.4 (Darriba *et al.,* 2012). The Markov Chain Monte Carlo (MCMC) simulations were run for 600 000 generations until the standard deviation of the split frequencies reached 0.01. All PSRF+ values reached the value of 1. The tree was finally modified and edited in FigTree and Adobe Illustrator. Standard estimators of diversity were calculated using DnaSP (Librado and Rozas, 2009).

Supplementary note S4: Double-digest restriction-site associated DNA sequencing (ddRAD-seq) library preparation

A double-digest RAD-sequencing protocol was carried out on DNA from the 61 AMF isolates, with three to five independent biological replicates of each isolate. Between 1 to 25ng/µl of genomic DNA, obtained from *in vitro* plates of each isolate, was digested with *Eco*RI and *Mse*I enzymes (NEB). The restriction digestion mix (RDM) contained 0.9 µl 10 X T4 DNA ligase buffer (NEB), 0.15M of NaCl, 0.15mg/ml of BSA (NEB), 1 unit of MseI (NEB) 10 000 U/ml, 5 units of EcoRI (NEB) 20 000 U/ml, and 0.85 µl of ddH_2O. Three µl of RDM were mixed with 6µl of genomic DNA at an ideal concentration of 25 ng/µl. The digestion was carried out in a Biometra thermocycler. Incubation time was 2 hours at 37°C and inactivation of the enzyme was at 65°C for 20 minutes.

Ligation of individual barcodes was carried out by adding 1µl (1µM) of EcoRI-P1 adapter, containing individual barcodes, to 9µl of digested DNA and RDM and completed with 1.6µl of a ligation mix containing 1X of T4 DNA ligase buffer (NEB), 0.08 M of NaCl, 0.08 mg/ml of BSA (NEB), 6.25µM of MseI-P2 adapter and 335 units of T4 DNA ligase (NEB) 2 000

000 U/ml. The volume was then completed with 0.0125 µl of ddH_2O. The ligation was incubated for 6 hours at 16°C followed by heat-inactivation for 10 minutes at 65°C. The final volume was then completed to 65 µl with ddH_2O. This step was followed by an AMPure (Agencourt AMPure XP beads, Beckman and Coulter) purification step with a 1X volume of AMPure beads and a PCR step to ligate Illumina primers. PCRs were repeated twice for each sample and then pooled. A second boost PCR was run in order to reduce primer dimer and heteroduplex formation. Details of these two PCR protocols are as follows: PCR master mix for amplification of ddRAD-seq fragments were run in a total volume of 20 µl, with 4µl of purified DNA after AMPure bead purification and 16µl of master mix. The master mix of 16 µl contained 1.25X of Q5 HF buffer (NEB), 0.78 mM of dNTP, 0.42 µM of each of the Illumina PCR primer, 1.25X of High GC enhancer (NEB), 0.4 units of Q5 HF polymerase 2 000U/ ml (NEB) finally the volume was completed with 6.3 µl of ddH_2O. The PCR reaction was run on a Biometra T1 Thermocycler. The thermocycler was programmed as follows for the amplification: 98°C for 30s, 18 cycles of amplification at 98°C for 20 sec, 60°C for 30 sec and 72°C for 40 sec, a final elongation of 10 min at 72°C.

A second PCR step was run in order to reduce primer dimer and heteroduplex formation. A 2 µl aliquot of a master mix was added to each previous PCR of each sample. Two µl of master mix were composed of 1X Q5 HF buffer (NEB), 2 mM of dNTP, 3.35 µM of each of the Illumina PCR primers. The master mix was then completed to 2 µl with 0.1 µl of ddH_2O. Amplification was run on a Biometra T1 Thermocycler at 98°C for 3min, 60°C for 2 min and finally 72° for 12min. The PCR product was then run on a 1.5% agarose gel for 30-40min at 120V.

Between 36 to 90 samples were pooled in each of the four libraries. The libraries were then precipitated and purified with AMPure beads to remove adapter dimers. Library quality was finally checked on a Fragment Analyser. Each library was then sequenced in one lane as paired-end reads on a HiSeq Illumina sequencer at the Lausanne Genomics Technologies Facility (LGTF).

Supplementary note S5: Biological replication

Biological replicates were obtained by splitting a culture of an isolate into between ten and twenty parts, which were then sub-cultured on new plates. These were then grown for three months. Three to five spore-producing plates of each isolate were subjected to ddRAD-seq as

independent biological replicates. One plate corresponded to one extraction and one sequenced biological replicate. The ten Swiss isolates sequenced in this study, as well as the 20 isolates sequenced by Wyss *et al.*, 2016, underwent a slightly different protocol, where spores were extracted together from 3 plates and represented one sequenced biological replicate. In the study of Wyss et *al.*, 2016, there were 3 replicates per isolate.

Supplementary note S6: Trimming and demultiplexing of reads (ddRAD-seq)

Illumina adapters were first removed using TagCleaner 0.14 (Schmieder *et al.*, 2010) followed by a quality-trimming step with PrinSeq-lite 0.20.4 lite (Schmieder *et al.*, 2011). Reads containing uncalled bases (N) or shorter than 50 bp were removed. The de-multiplexing of the reads of each individual barcode was processed with *process_radtags* from the Stacks pipeline 1.21 (Catchen *et al.*, 2011). The coverage calculation for each sample was made using Samtools 0.1.19 (Li *et al.*, 2009).

Supplementary note S7: BP&P analysis A10, species delimitation under the multispecies coalescent (MSC) model.

The BPandP model chosen for this analyses was A10 (BPP manual and Yang and Rannala, 2010; Rannala and Yang, 2013), using a reversible-jump MCMC species delimitation approach that ran for 500 000 generations and with a burn-in period of 50 000 generations and a fixed species tree. Three scenarios were tested by varying two priors; population size (θ) and root age (τ0) following Leaché and Fujita (2010). A first run was set up with a small population size (θ [2,2000]) and small divergence time (τ0, [2,2000]). In the second run, a large ancestral population size (θ [1,10]) and a deep divergence time (τ0, [1,10]) was used. Finally, the last prior values were established with a large ancestral population and a shallow divergence time (θ [1,10], τ0, [2,2000]). Results are presented in the BP&P tree in Figure 6. A speciation probability of > 0.95 is considered as high and confident for a speciation event. Every analysis was run twice and the ESS values were verified as being above 1000 for all parameters.

Supplementary note S8: Data quality assessment

The sequence data appeared to be good quality. There are four main reasons leading to confidence in the sequence quality for answering the main questions. First, holotypes of two main species clearly defined two main clusters of species for *R. irregularis* and *R. intraradices*, based on the ddRAD-sequence data, as well as sequences of the PTG marker.

Secondly, distances among 29 individuals of two different genotypic groups of *R. irregularis*, calculated from the ddRAD-seq markers, were strongly correlated with Jaccard distance based on previously published microsatellite data from Croll *et al.*, 2008. Third, the different biological replicates, sequenced with independent barcodes and in independent PCR reactions, clustered strongly together (Supplementary Figure S1), suggesting no strong bias due to PCR error, sequencing error or incorrect genetic distance analysis. Finally, Wyss et *al.*, 2016 showed that this ddRAD-seq protocol and pipeline were reliable for studying inter-isolate genetic diversity in *R. irregularis*.

Supplementary note S9: Measurement of hyphal density

One day before DNA extraction, each *in vitro* plate was photographed. Eight photographs of each plate were taken, always using the same light and magnification. After sequencing and identification of the different genetic groups, nine isolates of three genetic groups defined during study (Gp1, Gp3 and GP4) and that were taken in picture were measure for hyphal density. In each picture, three random vertical transects were drawn and numbers of intersections of mycelium with the line were recorded.

This resulted in 120 transects per isolate, except for DAOM197198-CZ, which resulted in 72 transects. Using R software (R core Team 2014, version 3.3.2) we then tested whether hyphal density significantly differed among the three genotypic groups. A mixed model was implemented with the function *lmer* of the lme4 package (Bates et *al.*, 2014). The genetic groups were considered as a fixed effect, the niched isolates in the genetic groups as well as the plates niched in each isolate were taken as the first random effect and were included in the model. The second random effect component of the model was added as the isolates niched into the genetic groups.

Chapter 3

Within-species phylogenetic relatedness of a common mycorrhizal fungus predicts plant community structure

Romain Savary[*], Lucas Villard[*] and Ian R. Sanders

Department of Ecology and Evolution, Biophore Building, University of Lausanne, 1015 Lausanne, Switzerland

[*] These authors contributed equally to this work.

This chapter was submitted to the Journal of Ecology, August 11, 2017

Abstract

Arbuscular mycorrhizal fungi (AMF) have been shown to influence plant community structure and diversity. Studies based on single plant – single AMF isolate experiments show that within AMF species differences lead to large differential growth responses of different plant species. Because of these differential effects, genetic differences among isolates of an AMF species could potentially have strong effects on the structure of plant communities.

In this study, we tested the hypothesis that within species variation in the AMF *R. irregularis* significantly affects plant community structure and plant co-existence. We took advantage of a recent genetic characterization of several isolates of this species using double-digest restriction-site associated DNA sequencing (ddRADseq). This allowed us to test not only for the impact of within AMF species variation on plant community structure but also for the role of the *R. irregularis* phylogeny on plant community metrics. Nine isolates of *R. irregularis,* belonging to three different genetic groups (Gp1, Gp3 and Gp4), were used as either single inoculum or as mixed diversity inoculum. Plants in a mesocosm representing common species that naturally co-exist in European grasslands were inoculated with the different AMF treatments.

We found that within-species differences in *R. irregularis* did not strongly influence the performance of individual plants or the structure of the overall plant community. However, the equilibrium of the plant community was affected by the phylogeny of the fungal isolates, where more closely-related AMF isolates were more likely to affect pant community evenness in a similar way compared to more genetically distant isolates.

Synthesis. This study is the first to date that underlines the effect of within AMF species variability on plant community structure. While differential effects of the AMF isolates were not strong, it is surprising that within a single AMF species enough functional variability exists that can change the equilibrium of a plant community and that this linked to the evolutionary history of the AMF species.

Keywords: AMF, *Rhizophagus irregularis*, plant community, plant-plant interactions, phylosignal, evenness

Introduction

Soil microorganisms that influence plant-plant interactions play a central role in terrestrial ecosystems (Wardle *et* al., 2004). This is particularly true for arbuscular mycorrhizal fungi (AMF; phylum Glomeromycota), which are considered the commonest of plant root symbionts, due to their unique capacity to form endosymbioses and to exchange nutrients with 74% of land plants (van der Heijden, Martin, Selosse, & Sanders, 2015). During the last decades, various beneficial effects of these fungi on different plant species were reported such as an increase in plant growth and plant nutrient acquisition, (van der Heijden et al., 1998), greater resistance to pathogens (Pozo & Azcón-Aguilar, 2007) and herbivores (Koricheva, Gange, & Jones, 2009), and increasing tolerance to drought, high salinity and pollutants (Al-Karaki, McMichael, & Zak, 2004; Hajiboland, Aliasgharzadeh, Laiegh, & Poschenrieder, 2010; Gonzalez-Chavez, Harris, Dodd, & Meharg, 2002). In addition to direct effects on plant physiology, AMF have also been shown to alter competitive interactions between plants (van der Heijden et al., 1998). Consequently, this impacts common metrics of plant community structure, such as community richness, community evenness (i.e the relative abundance of community members) and community productivity. For example, removing AMF from nutrient-poor tallgrass prairies, where the dominant plant is highly mycotrophic, had a tendency to favor facultative mycotrophic plants, thus increasing total community evenness and richness without increasing total productivity (Hartnett & Wilson, 1999). Conversely, in grasslands, dominated by grass species that derive little benefit from the association, AMF have been observed to favour the productivity of subordinate forbs (van der Heijden et al., 1998), thus increasing community evenness and richness (Mariotte et al., 2013). Other studies also demonstrated that AMF could promote or limit community productivity depending on the AMF taxon involved, regardless of plant community species richness (Klironomos, McCune, Hart, & Neville, 2000).

In experiments studying the effect of an AMF species inoculated on a single plant species, the range of growth responses of plants was found to vary greatly depending on the combination of plant species and AMF species tested; spanning a range of response from highly positive to negative (Klironomos, 2003). One possible explanation for this effect was that AMF species were not equivalent in their functional characteristics (Hart & Reader, 2002) and, thus, impacted the outcome of a specific AMF-plant interaction. Later on, Powell et al. (2009) showed that some of these differences were correlated with AMF phylogeny. Indeed, they

found that some fungal quantitative traits of members of the Glomeromycota phylum appear to be phylogenetically conserved at the family level. Consequently, this created a similar conservatism towards plant response. Similarly, it was shown that the plant response varied greatly while in symbiosis with different genotypes of a single AMF species. (Munkvold, Kjoller, Vestberg, Rosendahl, & Jakobsen, 2004; Mensah et al., 2015; Koch, Croll, & Sanders, 2006; Koch, Antunes, Maherali, Hart, & Klironomos, 2017). Depending on the species tested, the amplitude of variation in plant response due to within-AMF species differences was similar, or even higher, than the variation in effects among AMF species. These observations lead Sanders and Rodriguez (2016) to suggest that there is an inconsistency between the level of AMF phylogenetic resolution used by experimental ecologists and the level of functionality in the AMF phylogeny. Indeed, most ecological studies to date used one representative isolate for each species (van der Heijden et al., 1998; Vogelsang, Reynolds, & Bever, 2006; Maherali & Klironomos, 2007), making the assumption that the within-AMF species genetic diversity and its plant effect was homogenous across any isolate of that species. In the light of the results of Munkvold, Kjoller, Vestberg, Rosendahl, and Jakobsen, 2004; Mensah et al., 2015; Koch, Antunes, Maherali, Hart, and Klironomos, 2017, this assumption would seem to be compromised. From the observed effects of high variation in a plants response to different genotypes of a single AMF species, we could expect that different *R. irregularis* genotypes might differentially affect plant-plant interactions in a plant community. Moreover, according to the findings of Powell et al. (2009), we could also expect that intra-specific variation in fungal traits, plant response and community response might be explained by phylogenetic conservatism. Testing intraspecific phylogenetic effects on plant communities were previously difficult because of the lack of accurate data on intraspecific genomic differences in AMF.

According to recent estimates, the Glomeromycota phylum is composed of 300 to 1600 AMF species (van der Heijden, Martin, Selosse, & Sanders, 2015) and the vast majority of ecosystems harbour an assembly of several AMF species. Studies that manipulated AMF species richness in a plant community have shown that there was a positive relationship between AMF diversity and community productivity (van der Heijden et al., 1998). This effect could be the result of a functional complementarity of co-occuring taxa, since Maherali and Klironomos (2007) have shown that more phylogenetically diverse AMF communities were more stable and resulted in greater plant biomass compared to the effects of AMF communities composed of taxa that were phylogentically similar. In the putative framework

of a phylogenetic conservatism of AMF traits and plant responses, one could expect that AMF communities that are phylogenetically diverse at the intraspecific level will positively impact the different community metrics, as it has been suggested that a mix of different genotypes could affect ecosystem function as much as a mix of species (Johnson, Martin, Cairney, & Anderson 2012).

In this study, we took advantage of a recent characterization of intraspecific genetic variability in one of the commonest AMF species *Rhizophagus irregularis,* using double-digest restriction-site associated DNA sequencing (ddRADseq) (Savary et al., submitted). We also chose to use this AMF species because of the previously documented variability in quantitative traits and effects on plant growth (Koch, Croll, & Sanders, 2006). We tested the importance of intraspecific diversity of *R. irregularis* on different plant species co-existing within a plant community and on characteristics of plant communities. We sought to determine if genetic relatedness among isolates of *R. irregularis* was correlated to the putative effects on single plant species within a plant community and on characteristics of mesocosm plant communities. We also wanted to test if increasing phylogenetic diversity in a community of *R. irregularis* could positively affect plant community productivity and evenness.

We built artificial mesocosms with a fixed plant community composed of six plant species, that are typical of central European calcareous grasslands. We chose nine *R. irregularis* isolates equally distributed into three genetic groups describing a large part of the known *R. irregularis* intraspecific genetic variation (Savary et al., submitted). The genetic groups are named Gp1, Gp3 and Gp4 (Savary et al. submitted). Both Gp3 and Gp4 naturally co-exist in several locations and all three genetic groups occur in central Europe (Savary et al., submitted). All isolates were cultured for several generations in an *in vitro* culture system in order to remove potential environmental effects from their place of origin that could otherwise be confounded with genetic effects. These isolates were either inoculated singly into the mesocosms or as a combination of isolates. This design allowed us to test three main questions: i) Do different isolates of *R. irregularis*, or combinations of the isolates, differentially colonize plants and differentially affect plant responsiveness, community structure and productivity? ii) Do individual plant species of a community respond more similarly to more closely genetically related AMF isolates compared to more distant ones? iii) Is the mycorrhizal responsiveness and structure of a plant community more similar in response to more closely genetically related AMF isolates compared to more distant isolates?

Material & Methods

Soil and fungal inoculum

A natural clay soil collected from a calcareous grassland at the University of Lausanne, Lausanne, Switzerland (46°31'32.0016" N, 006°34'46.6068" E) was used for this experiment. The soil was sieved through a 1 cm mesh, then mixed with quartz sand in a ratio 2:3. This mixture was then steam-sterilized at 105°C for three consecutive sessions of 20 min. Round 4.6 litre pots (30 cm diameter) were filled with the sterilized soil:sand mixture.

Nine *R. irregularis* isolates were chosen in three different genetic groups (Gp1, Gp3 and Gp4), representing the major part of the diversity described in this species (Savary et al., submitted). These were: Gp1 - 5B, ESQLS69, LPA54; Gp3 - A4, DAOM229457, DAOM240409; Gp4 - A1, DAOM197198-CZ, DAOM240159 (Fig. 1a). The delineation of this variation was based on previous ddRADseq data generated from DNA of 59 isolates of *R. irregularis* isolated from several geographical locations and comprising 6888 sites in the genome where single nucleotide polymorphisms (SNPs) occurred. Each of the 9 isolates was first grown in an identical *in vitro* environment for 5 months (Bécard & Fortin, 1988). In order to create a natural soil inoculum for each isolate, following Wagg, Jansa, Stadler, Schmid, and van der Heijden (2011b), we inoculated *P. lanceolata* with 500 *in vitro*-produced spores in ten 0.5 litre pots in a sterile autoclaved Oil-Dri® granular clay (Oil Dri corporation of America) with quartz sand in a 3:2 ratio. Ten *P. lanceolata* were mock inoculated with water in order to produce a substrate representing a mock-inoculated treatment. The *P. lanceolata* plants were grown for five months and then the soil of the ten pots per isolate were mixed to create the 9 inoculum stocks and substrate for the mock-inoculated treatment.

Plant material

In order to construct simple artificial European calcareous grassland community in the greenhouse, we chose six plant species commonly found in this type of phytosociological association (classified as *Arrhenatherion* according to Delarze & Gonseth, 2008). The following plant species were chosen to be equally partitioned across the different functional groups (Grime, Mackey, Hillier, & Read 1987). Thus, we chose two legumes, *Trifolium pratense* and *Lotus corniculatus*, two grasses, *Arrhenatherum elatius* and *Festuca pratensis* and two forbs *Prunella vulgaris* and *Knautia arvensis*. All these species are known to form the arbuscular mycorrhizal symbiosis (Wang & Qiu, 2006). Seeds of these plants were

obtained from UFA Samen (Switzerland) and germinated in trays on the same substrate used for the experiment. After two weeks of growth, one individual of each species was planted in a circle in each pot (Fig. 1b). The position was randomly chosen except that plants of the same functional group were never planted next to each other but always at the most distant location. Thus, *Arrhenatherum elatius* was always the most distant plant from *Festuca pratensis, Trifolium pratense* was always the most distant plant from *Lotus corniculatus* and *Prunella vulgaris* was always the most distant plant from *Knautia arvensis* (Fig 1b).

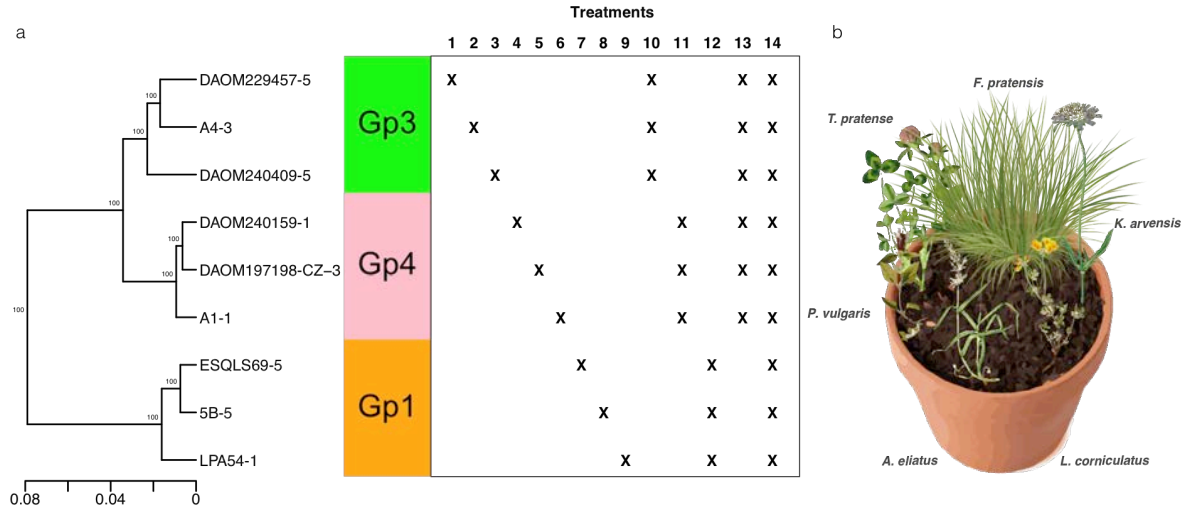


Figure 1 Study design (a) Phylogeny of *R. irregularis* isolates from three genetic groups used as treatments either as single inoculants (treatments 1-9) or as mixed inoculant (treatments 10-14). **(b)** Representation of one mesocosm with the six plants.

Fungal treatments and experimental design

Fourteen mycorrhizal treatments and one non-mycorrhizal (NM) treatment were applied to the plant communities and were replicated ten times (Fig. 1a). We inoculated each pot of the 9 single inoculation treatments (treatments 1-9) with 100g of inoculum. Treatments with a co-inoculation of three isolates (Gp1, Gp3 and Gp4; treatments 10-12) received 33.3g of inoculum of each isolate. In the case of Gp3+Gp4 (treatment 13) we used 16.6g of inoculum of each of the six isolates. Finally in the co-inoculation treatment with all isolates (treatment 14), 11.1g of inoculum of each isolate was added per pot. The fifteen treatments were randomly arranged on a table in the greenhouse. This procedure was repeated 10 times for the 10 replicates separated on 10 tables. Tables were regularly randomized to avoid microclimate effects. The greenhouse conditions were constant at, 24°C, 60% RH and 12h of daylight. The communities were grown for 3.5 months from May 4th 2015 to the August 19th 2015.

Harvest and measurements

After 105 days of growth, the shoots of each plant in the community were harvested and dried at 70°C for three days and weighed to obtain the above ground dry mass (ADM) for each of the 900 plants. A small part of the roots of each plant was collected and frozen at -20°C for later measurement of AMF colonization by staining. Roots were stained and AMF colonization was measured (see supplementary note S1). The roots from each mesocosm of all the six plants were washed dried together and weighed, thus, giving the total community root dry mass (RDM). The sum of all plant ADM from one pot plus the RDM of this pot resulted in total dry mass (TDM) of the community. Inflorescence number was counted at 78 and 105 days of growth.

Statistical analysis

In order to have a standardized measure of mycorrhizal effects on the different plant species, plant individual responsiveness was calculated following Gange and Ayres (1999) based on the mean ADM for each plant species in the non-mycorrhizal treatment. We used the data collected for each individual in a mesocosm to build the metrics of community structure. TDM was considered as the community productivity. Mean community responsiveness and mean AMF colonization were obtained by averaging the individual responsiveness across plant species and AMF colonization across plant species, respectively. Community evenness was measured with Pielou's evenness index (Pielou, 1975) using the diversity function in vegan 2.3-3 (Oksanen et al., 2016). This was calculated by dividing the Shannon index of ADM by the log of the number of individuals in the community. Evenness represents a value of equality (1) or inequality (0) of biomass partitioning among species within a mesocosm. The AMF colonization evenness was calculated in the same way in order to assess equality or inequality of AMF colonization among the six plant species within a mesocosm. This could be considered as a proxy for AMF preferences within a mesocosm. Significant correlations between mesocosm variables, as well as the relationships between single plant data and mesocosm averages, were assessed using Pearson's product moment and polynomial regressions.

The effect of plant species and mycorrhizal treatments on AMF colonization, plant responsiveness and number of flowers produced were analyzed using a two-way ANOVA. Significant pairwise comparisons were assessed using a Tukey HSD post-hoc test. Significant

differences among AMF treatments towards plant community metrics were assessed using a one-way ANOVA and a Tukey HSD post-hoc test.

The R. irregularis *phylogeny*

All nine isolates used in this study were previously sequenced with ddRAD-seq with a minimum of 3 replicates (Savary et al., submitted). The raw sequence reads were retrieved from the NCBI bioproject (accession number PRJNA326895) and were trimmed and analysed following the workflow of Wyss, Masclaux, Rosikiewicz, Pagni, and Sanders (2016) and Savary et al. (submitted). Genetic distance matrices among the nine isolates were calculated based on the scalar distance method of Wyss, Masclaux, Rosikiewicz, Pagni, and Sanders (2016). For this calculation, data used were taken from the replicate with the deepest sequencing of each of the nine isolates and with the highest number of markers available.

Phylogenetic signal

The package phylosignal (Keck, Rimet, Bouchez, & Franc, 2016) was used to test phylogenetic signals between the phylogeny of the 9 isolates and i) AMF colonization and mycorrhizal responsiveness across the 6 plant species and ii) the mean AMF colonization, mean mycorrhizal responsiveness, evenness, total productivity and mean flower production after 78 and 105 days per mesocosm.

Five phylogenetic signal indicators were calculated on variables of each of the plant species and on community metrics of the mesocosms. These were Moran's I index (Moran, 1948, 1950), Abouheif's Cmean index (Abouheif, 1999), Blomberg's K and K* (Blomberg, Garland, & Ives, 2003) and Pagel's λ (Pagel, 1999). These were then tested against the null hypothesis of a random trait with no significant signal (Keck, Rimet, Bouchez, & Franc, 2016). Plant and mesocosm variables showing significant phylogenetic signal towards AMF isolates were kept for following analyses on co-inoculation treatments. AMF community phylogenetic diversity for treatments involving more than one AMF taxa was calculated using Faith's PD (Faith, 1992) and relationships between the retained variables and phylogenetic diversity were assessed using quantile polynomial regression on the median, 20th and 80th quantiles of the response variable.

Results

Effect of fungal treatments on individual plant species

Overall, plants in the fourteen treatments exhibited high levels of AMF colonization in all six plants with a global mean of 67.6% ± 19.7% (SD). Fifty-seven plants in the NM treatment exhibited no AMF colonization. Three plants were measured with a very small amount of AMF colonization (less than 5%). These were considered as a contaminant. Colonization of the roots by the fungi was significantly affected by the different AMF treatments as well as by plant species identity. However, the interaction between these two variables was not significant (Table 1a). Isolates A4 (Gp3) and ESQLS69 (Gp1) were significantly the lowest colonizers and the three isolates of Gp4: A1, DAOM240159 and DAOM197198-CZ were the highest colonizers (Fig. 2a). *R. irregularis,* independent of isolate identity, colonized a significantly greater proportion of the roots of *K. arvensis*, *P. vulgaris* and *T. pratense* than the dominant plant, *F. pratensis* (Fig. 2b).

Large and significant differences in mycorrhizal responsiveness occurred among the different plant species but was not significantly influenced by the different AMF isolates (Table 1a). Mycorrhizal responsiveness of the plant community was positive (23.13 ± 107.82 SD) in the majority of fungal treatments (Fig. 2c). Mycorrhizal responsiveness of the grasses *A. eliatus* and *F. pratensis*, and *L. corniculatus* was negative and differed significantly from the positive mycorrhizal responsiveness of *K. arvensis*, *P. vulgaris* and *T. pratense* (Fig. 2d). *F. pratensis* responsiveness decreased with increasing AMF colonization (cor= -0.21, p= 0.013 ; Fig. S1). In contrast, mycorrhizal responsiveness of *P. vulgaris* and *T. pratense* increased with increasing AMF colonization (cor= 0.30, p<0.001, cor= 0.25, p= 0.004, Fig. S1). There was no significant correlation between mycorrhizal responsiveness and AMF colonization in the remaining plant species.

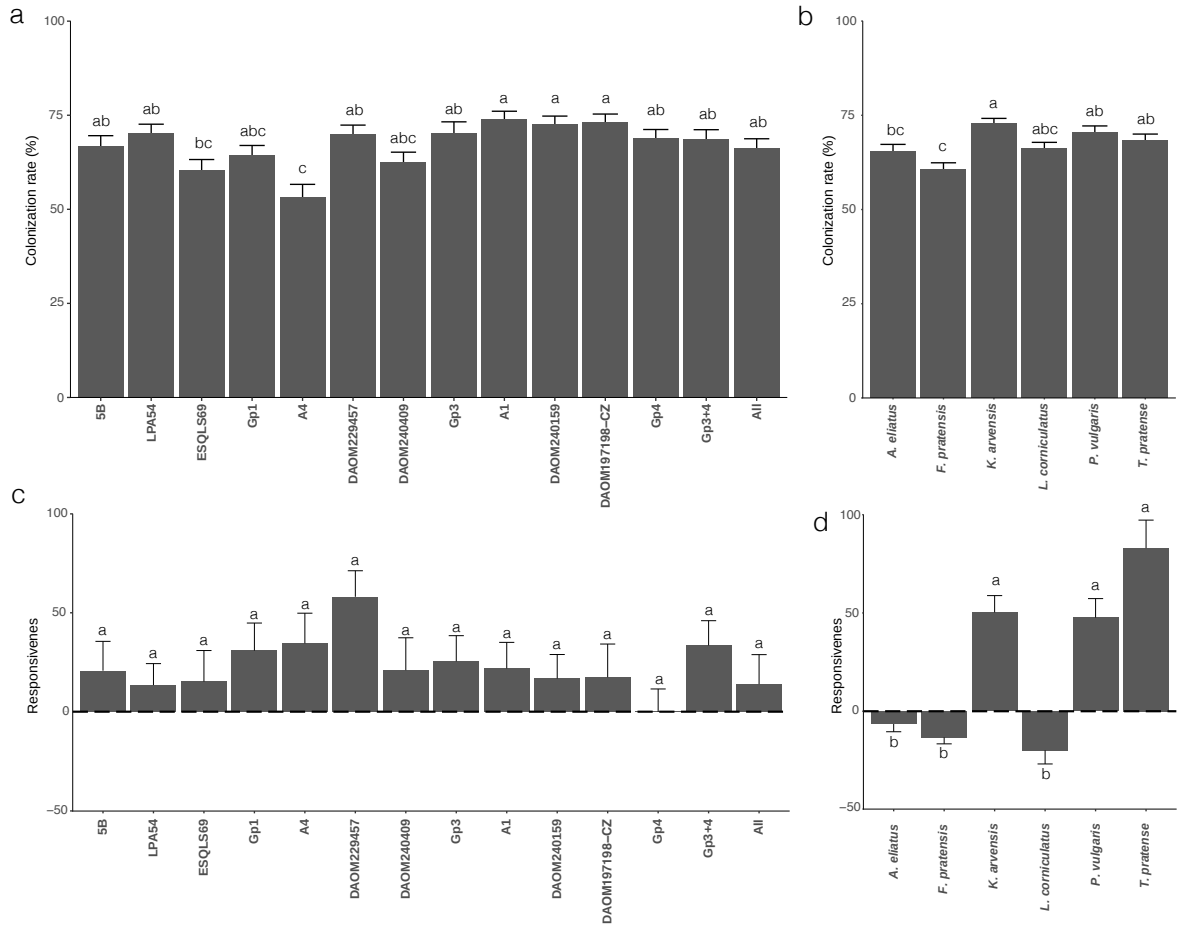


Figure 2 AMF colonization and responsiveness in the different treatments and different plant species

(a) Mean and the standard error of AMF colonization (% root length) according in the 14 treatments, and **(b)** in the six plants species. **(c)** Mean and the standard error of the responsiveness of each plant in the different treatments and **(d)** in the six different plant species of the community.

Effect of fungal treatments on plant communities

The dry mass production of the mesocosms either as RDM, ADM or as total productivity (TDM), was not significantly affected by any of the AMF treatments (Fig. S2a and S2b, Table 1b). Plant community evenness and AMF colonization evenness were not significantly different across all treatments (Fig. S2c; Table 1b). However, a number of interesting correlations were observed between variables. Mean AMF colonization per mesocosm was significantly and positively correlated with plant community evenness (cor=0.256, p=0.0023, Fig. S3), as was mean AMF colonization per mesocosm with AMF colonization evenness (cor=0.724, p<0.001, Fig. S4). Mean mesocosm responsiveness was significantly and positively correlated with mean AMF colonization (cor=0.300, p<0.001, Fig. 3a) and total productivity (cor=0.481, p<0.001, Fig. 3b). Plant community evenness was significantly and

positively correlated with mean mycorrhizal responsiveness (cor=0.752, p<0.001, Fig. 3c). The increase in plant community evenness was associated with a strong significant decrease in the relative contribution of *F. pratensis* to the total above-ground mesocosm productivity and a significant increase in the relative contribution of all the other plants (Table S1; Fig. 3d).

Table 1 Results of ANOVA on (a) AMF colonization, mycorrhizal responsiveness and flower production at 85 and 105 days per plant; **(b)** above-ground dry mass evenness, root dry mass, above-ground dry mass, total productivity, flower production at 85 and 105 days, and AMDF colonization evenness per mesocosm.

Source of Variation	*df*	*MS*	*F*	
a. Variables measured on individuals plants				
Colonization rate (%)				
AMF	13	1831.27	3.812	***
Plant	5	2450	6.491	***
AMF x Plant	65	328.77	0.871	
Residuals	715	377.43		
Mycorrhizal responsiveness				
AMF	13	10802	1.041	
Plant	5	241361	23.262	***
AMF x Plant	65	12276	1.12	
Residuals	724	10954		
Flower production 85 days				
AMF	13	17.11	0.5423	
Plant	5	605.85	32.211	***
AMF x Plant	65	18.79	0.999	
Residuals	756	18.81		
Flower production 105 days				
AMF	13	6.605	1.708	
Plant	5	60.545	15.658	***
AMF x Plant	65	4.114	1.064	
Residuals	756	3.867		

b. Mesocosm variables					
ADM Evenness					
	AMF	13	0.00193	1.03	NS
	Residuals	126	0.0019		
Root drymass					
	AMF	13	6.861	0.447	NS
	Residuals	126	1931.62		
Above-ground drymass					
	AMF	13	323.49	0.671	NS
	Residuals	126	35		
Total productivity					
	AMF	13	44.2	0.704	NS
	Residuals	126	62.78		
Flower production 85 days					
	AMF	13	39.63	1.6778	NS
	Residuals	126	23.62		
Flower production 105 days					
	AMF	13	102.68	0.866	NS
	Residuals	126	118.58		
Colonization evenness					
	AMF	13	0.00197	1.581	NS
	Residuals	116	0.00125		

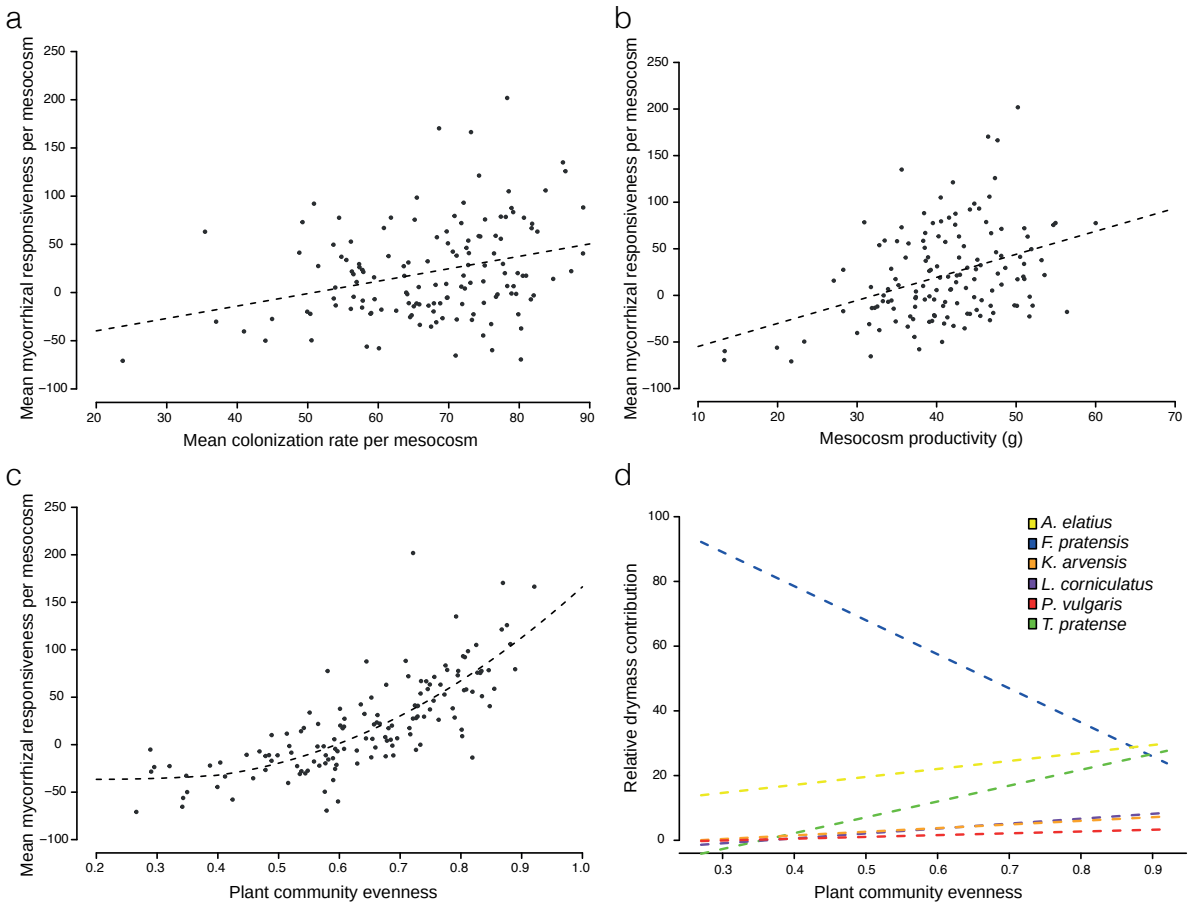


Figure 3 Community metrics and relative dry mass contribution of each plant to the community

(a) Correlation between the mean mycorrhizal responsiveness and the mean colonization rate (cor=0.300, p<0.001) **(b)** Correlation between the mean plant responsiveness and the mesocosm productivity (TDM, cor=0.481, p<0.001) **(c)** Quadratic relation between the mean plant responsiveness and the plant community evenness per mesocom (cor=0.752, p<0.001). **(d)** Relative dry mass contribution of each plant species to the mesocosm according to the estimator of plant community evenness. All models were significant (p<0.001)

Phylogeny of the nine isolates

Out of the ddRAD-seq data (Savary *et al*. submitted) from the nine isolates used in this study, we retrieved the data from the biological replicate of each isolate that had been sequenced the deepest and with the highest SNP calling quality. We were able to retrieve 60668 shared sequence positions, which contained 16311 SNPs, 1137 insertions, 1453 deletions and 11 MNPs. The much higher number of variable positions, in comparison to the study of Savary et al. (submitted), was due to the low number of isolates and replicates used. This enabled us to retrieve shared positions among all isolates that would not be shared among all isolates in a larger data set. This variation was used to build 100 genetic distance matrices from 5000 randomly chosen sites, following the method described in Wyss, Masclaux, Rosikiewicz, Pagni, and Sanders (2016). All nodes showed a support value of 100, clearly separating the nine isolates into three main groups (Fig. 1a). This tree was then used for phylogenetic signal analyses.

Testing for phylogenetic signals with plant species data

There was a significant within-species AMF phylogenetic signal in AMF colonization of *F. pratensis* in two out of the five tests (Fig. 4a), with a globally highest colonization by Gp4 isolates. Similarly, there was a significant within-species AMF phylogenetic signal in the mycorrhizal responsiveness of *F. pratensis* (Fig. 4b) in three tests out of the five. There was also a significant within-species AMF phylogenetic signal in the mycorrhizal responsiveness of *K. arvensis* for one test out of five (Fig 4b).

Testing for phylogenetic signals in plant community metrics

At the community level, the overall mean AMF colonization, RDM, ADM, TDM and flower production did not reveal a significant within-species AMF phylogenetic signal (Fig. 4c). In contrast, there was a significant within-species AMF phylogenetic signal in mean mycorrhizal responsiveness of the community in four out of five tests (Fig. 4c). There was no detectable within-species AMF phylogenetic signal in AMF colonization evenness (Col-evenness) (Fig. 4d). There was a significant within-species AMF phylogenetic signal in terms of community evenness (ADM-evenness) in three tests out of five (Fig 4d). Phylogenetic diversity of the AMF community, calculated with Faith's index, showed a significant quadratic relationship with plant community evenness ($y=0.023x^{-0.088} + 0.195$, $p=0.0475$, Fig. 5), showing the highest community evenness values at an intermediate level of phylogenetic diversity.

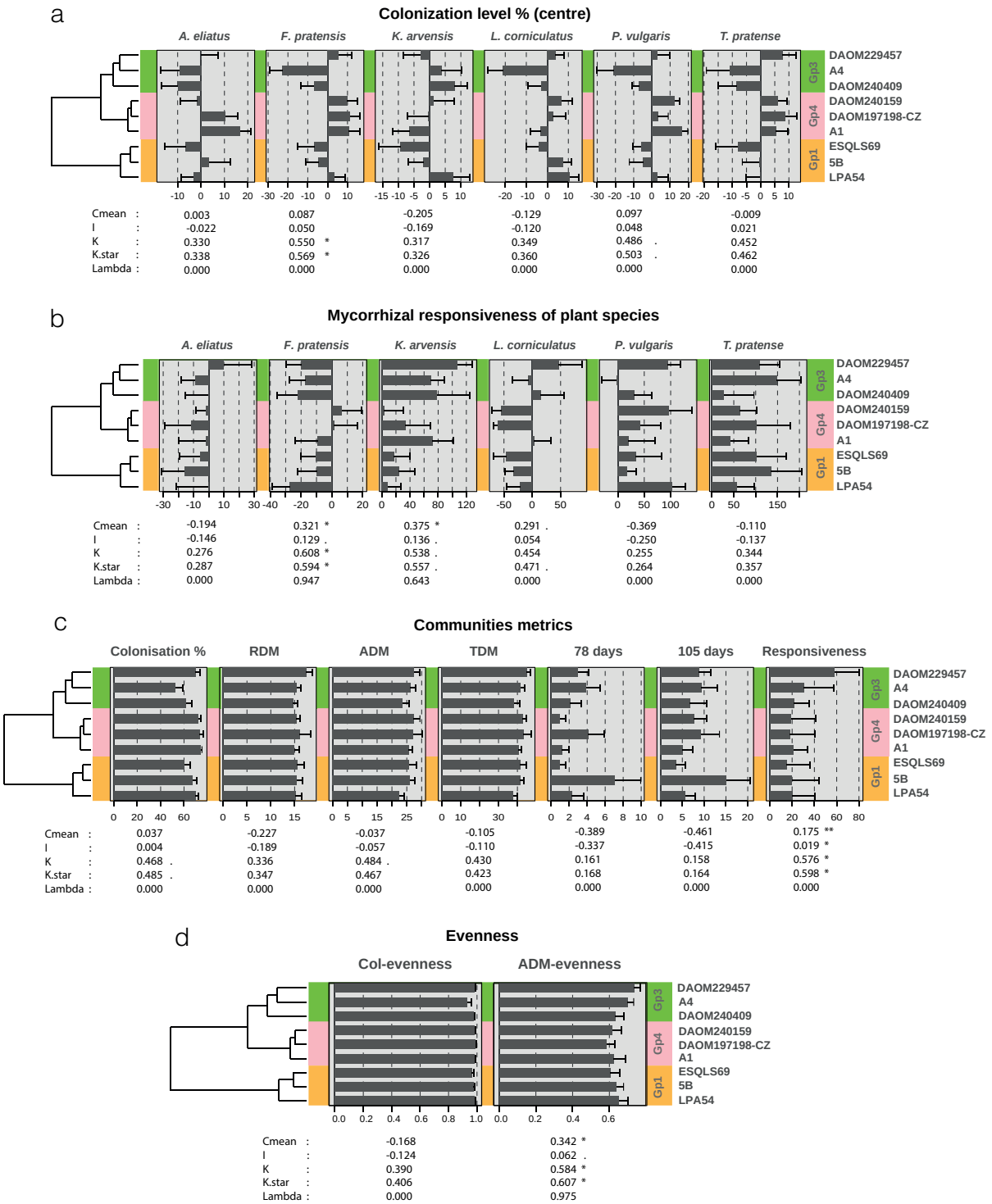


Figure 4 Individual plant species and mesocosm phylogenetic signals. Each of the phylogenetic signal plots are composed of the phylogeny of the nine isolates from the three genetic groups. The value of the different tests and their significance (. < 0.1, * < 0.05, ** < 0.01) are indicated under each set of bar plots. The phylosignals were calculated for each of the six plants independently for **(a)** the colonization level (%), here centred on the mean colonization per plant, and **(b)** for the responsiveness (not centered). Mesocosm phylosignals are presented **(c)** for mean colonization, RDM, ADM, TDM, flower production after 78 and 105 days and mean mesocosm responsiveness and **(d)** for mesocosm colonization repartition (Col-evenness) and plant community evenness (ADM-evenness)

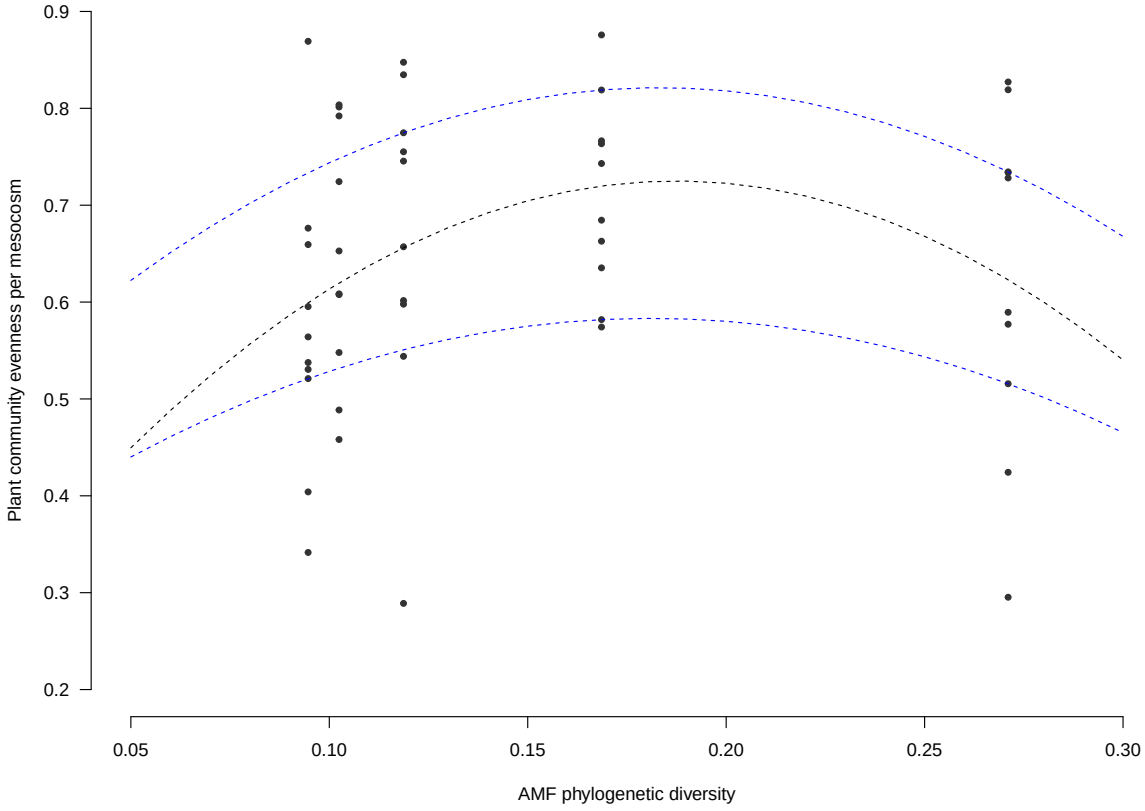


Figure 5 Relationship between plant community evenness and AMF phylogenetic diversity of the five treatments (10-14) with mixed AMF inoculant.

Discussion

Up to now, the interaction between AMF and plant species within a community was mostly investigated by manipulating AMF at the species level, rarely taking into account the phylogenetic relatedness of AMF isolates. To our knowledge, this study is the first to consider genetic differences among isolates of an AMF species as a source of potential functional trait variability that could differentially impact plant-plant interactions within a community. We show that despite the strong AMF intraspecific effect on single plant responses observed elsewhere (Munkvold, Kjoller, Vestberg, Rosendahl, & Jakobsen, 2004; Mensah et al., 2015; Koch, Antunes, Maherali, Hart, & Klironomos, 2017), the effect of *R. irregularis* isolates on each plant species and on the global community were weak. Though weak, the level of *R. irregularis* isolate relatedness did influence the response of some plant species in the community. Furtheormore, we observed a phylogenetic conservatism on plant community metrics of responsiveness and on community evenness.

i) Do different isolates of R. irregularis, or combinations of the isolates, differentially colonize plants and differentially affect plant responsiveness, community structure and productivity?

Although we observed significant differences in AMF colonization ability among the AMF isolates the overall mycorrhizal responsiveness of the community to these isolates did not differ significantly. This suggests that none of the *R. irregularis* strains used in this study was a general plant community growth promoter. Although the mycorrhizal responsiveness of the 6 plant species differed greatly, this was not influenced by the identity of the AMF isolates and, thus, there were no strongly observable differences in plant community structure and productivity due to the different AMF isolate treatments. Despite that number of previously published studies have shown differential effects of within-species AMF variation on the growth of different plant species, in single plant – single fungus experiments, our study did not support the hypothesis that such effects could lead to strong differences at the plant community level. Some of the plant species and AMF isolates used in this study had been shown in previous studies to differentially alter the growth of different plant species.

We found strong positive responsiveness of the subordinate plants of the community except in *L. corniculatus.* The two dominants plants, the grasses *F. pratensis* and *A. eliatus* responded negatively to being colonized by AMF. This resulted in fairly similar mesocosm productivity

in every AMF treatment. This observation might be explained by a limited amount of resources available in mesocosms leading to a saturation of the productivity. Such saturation has already been observed in similar experiments (van der Heijden et al., 1998). Therefore, the most relevant variation on plant community structure detectable in such an experimental design is the relative contribution of individual plants to community productivity. This is given by the evenness index, which describes the relative competition strength within the plant community (Mulder et al., 2004). Here, we did not find strong differences in evenness among treatments with different isolates. However, an increase in plant community evenness was mainly due to a decrease in *F. pratensis* drymass relative to an increase in *T. pratense* and *A.elatius*, as these three species contributed to almost 80% of the mesocosm productivity. This effect of dominance mediation by AMF has previously been observed in calcareous grasslands (Mariotte et al., 2013).

ii) Do individual plant species of a community respond more similarly to more closely genetically related AMF isolates compared to more distant ones?

When data collected for each plant species was analysed separately, a phylogenetic signal was detected on AMF colonization of *F. pratensis* meaning that the compatibility of *F. pratensis* and *R. irregularis* in a plant community is linked to the evolutionary history of the different isolates of *R. irregularis*. In addition, a significant phylogenetic signal in mycorrhizal responsiveness of *F. pratensis* as well as for *K. arvensis* was found. This suggests that the trait evolution of different *R. irregularis* isolates not only impacts the ability of the fungus to colonize a given plant species but also indicates that the outcome of the symbiosis in terms of plant growth is more likely to be similar of the AMF isolates are genetically closer. The potential preference of some *R. irregularis* genotypes for some plant species was previously observed (Croll et al., 2008) as similar *R. irregularis* genotypes tended to be more likely isolated from one plant species that was used as trap-culture than another plant species. This heterogeneity of preference among the plants could be explained by a slight co-adaptations between AMF genotypes and plant species and might favour diversity of AMF and allow the coexistence of closely related AMF genotypes within a community.

iii) Is the mycorrhizal responsiveness and structure of a plant community more similar in response to more closely genetically related AMF isolates compared to more distant isolates?

At the community level, mean mycorrhizal responsiveness of the community, as well as mean community evenness, were found to be significantly affected by the phylogenetic relationships of the *R. irregularis* isolates. These findings are coherent with the AMF phylogenetic effects observed on *F. pratensis* since community evenness was mainly associated with a decrease of relative ADM contribution of this plant species. *F. pratensis* was the strongly dominant plant of the mesocosm and, similarly to other grass species, *F. pratensis* is known to produce a large amount of roots compared to forbs or legumes (Grime, Mackey, Hillier, & Read 1987). Thus, we could expect that isolates of Gp4 would benefit more than other *R. irregularis* genotypes on dominant grasses. In contrast, members of Gp3 would more likely be associated with increased benefits on subordinate plants since this clade provoked a strong decrease in *F. pratensis* responsiveness, and concomitantly, an increase in *K.arvensis* and *T. pratense* responsiveness. The differences found among these phylogenetic groups suggest that some functional differences might exist at the intra-specific level in *R. irregularis.*

Increasing fungal taxonomic diversity by inoculating different mixes of several *R. irregularis* isolates on the same mesocosm did not significantly influence plant community productivity. Nevertheless, studies showing this relationship were mainly performed in field assays, where space and soil nutrient levels are less limited and, thus, saturation is not reached (van der Heijden et al., 1998). We found a significant quadratic relationship between AMF community phylogenetic diversity and plant community evenness, even though only five treatments in the experiment were effectively manipulating AMF phylogenetic diversity. Community evenness was the highest at intermediate level of AMF community phylogenetic diversity and dropped to a lower value where all isolates were combined. This observation suggests a saturating effect of the AMF community where the potential benefit of AMF on biomass partitioning within the plant community is no longer observable. Nevertheless, we did not test all the combinations of AMF isolates so the relationship observed might still be attributed to a complementarity effect occurring by chance or to the effect of one particularly good isolate included in the co-inoculations (Wagg, Jansa, Schmid, & van der Heijden, 2011a). However, evidence for coexistence of Gp3 and Gp4 in the same grassland has been assessed, but not for the other combinations of phylogroups (Savary et al., submitted). This favours the hypothesis of less niche overlap between these two phylogroups, as suggested by the phylogenetic signal, which therefore might increase functional diversity in this particular treatment (Maherali & Klironomos, 2007).

Perspectives and Conclusions

To our knowledge, these are the first results to show that within species diversity of these fungi and in particular phylogenetic relatedness can impact mycorrhizal responsiveness of dominant plants in a community and, consequently, biomass partitioning among a community of plants. These results extend the findings of Powell et al. (2009) in that phylogenetic conservatism of AMF functional traits on plant communities can exist within an AMF species and not only between major AMF clades. If confirmed, this feature is interesting because it suggests that the outcomes of plant-fungal and community-fungal interactions are genetically based and could be conserved over evolutionary time. Further studies should focus on AMF traits that are known to be of major importance to plant growth such as level of nutrient acquisition and transfer to the host and consider them in a phylogenetic context at AMF intra-specific level.

Acknowledgements

We thank Nicolas Ruch, Jérôme Wassef, Cynthia Meizoso, Noémie Gambino for their help in harvesting and Nicolo Tartini and Rafael Joss for their help in root staining. We also thank Aleš Látr and Miroslav Vosatka at Symbiom and Yolande Dalpé at GINCO for providing us with AMF isolates. Bioinformatics computations for the phylogenetic tree were performed at the Vital-IT (http://www.vital-it.ch) Center for High Performance Computing of the Swiss Institute of Bioinformatics. This study was funded by the Swiss National Science Foundation (Grant number: 31003A_162549 to IRS).

References

Abouheif, E. (1999) A method for testing the assumption of phylogenetic independence in comparative data. *Evolutionary Ecology Research,* **1,** 895-909.

Al-Karaki, G., McMichael, B. & Zak, J. (2004) Field response of wheat to arbuscular mycorrhizal fungi and drought stress. *Mycorrhiza,* **14,** 263-269.

Bécard, G., & Fortin, J.A. (1988) Early events of vesicular arbuscular mycorrhiza formation on Ri T-DNA transformed roots. *New Phytologist,* **108,** 211-218.

Blomberg, S.P., Garland, T., & Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution,* **57,** 717-745.

Croll, D., Wille, L., Gamper, H.A., Mathimaran, N., Lammers, P.J., Corradi, N. & Sanders, I.R. (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,** 672-687.

Delarze, R. & Gonseth, Y. (2008) *Guide des milieux naturels de Suisse.* Bussigny. 424 p. Rossolis.

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation,* **61,** 1-10.

Gange, A.C. & Ayres, R.L. (1999) On the relation between arbuscular mycorrhizal colonization and plant 'benefit'. *Oikos,* **87,** 615-621.

Gonzalez-Chavez, C., Harris, P.J., Dodd, J. & Meharg, A.A. (2002) Arbuscular mycorrhizal fungi confer enhanced arsenate resistance on *Holcus lanatus*. *New Phytologist,* **155,** 163-171.

Grime, J.P., Mackey, J.M.L., Hillier, S.H. & Read, D.J. (1987) Floristic diversity in a model system using experimental microcosms. *Nature,* **328,** 420-422.

Hajiboland, R., Aliasgharzadeh, N., Laiegh, S.F. & Poschenrieder, C. (2010) Colonization with arbuscular mycorrhizal fungi improves salinity tolerance of tomato (*Solanum lycopersicum* L.) plants. *Plant and Soil,* **331,** 313-327.

Hart, M.M. & Reader, R.J. (2002) Host plant benefit from association with arbuscular mycorrhizal fungi: variation due to differences in size of mycelium. *Biology and Fertility of Soils,* **36,** 357-366.

Hartnett, D.C. & Wilson, G.W.T. (1999) Mycorrhizae influence plant community structure and diversity in tallgrass prairie. *Ecology,* **80,** 1187-1195.

Johnson, D., Martin, F., Cairney, J.W.G. & Anderson, I.C. (2012) The importance of

individuals: Intraspecific diversity of mycorrhizal plants and fungi in ecosystems. *New Phytologist,* **194,** 614-628.

Keck, F., Rimet, F., Bouchez, A. & Franc, A. (2016) phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecology and Evolution,* **6,** 2774-2780.

Klironomos, J.N. (2003) Variation in plant response to native and exotic arbuscular mycorrhizal fungi. *Ecology,* **84,** 2292-2301.

Klironomos, J.N., McCune, J., Hart, M. & Neville, J. (2000) The influence of arbuscular mycorrhizae on the relationship between plant diversity and productivity. *Ecology Letters,* **3,** 137-141.

Koch, A.M., Antunes, P.M., Maherali, H., Hart, M.M. & Klironomos, J.N. (2017) Evolutionary asymmetry in the arbuscular mycorrhizal symbiosis: Conservatism in fungal morphology does not predict host plant growth. *New Phytologist,* **214,** 1330-1337.

Koch, A.M., Croll, D. & Sanders, I.R. (2006) Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. *Ecology Letters,* **9,** 103-110.

Koricheva, J., Gange, A.C. & Jones, T. (2009) Effects of mycorrhizal fungi on insect herbivores: A meta-analysis. *Ecology*, **90,** 2088-2097.

Maherali, H. & Klironomos, J.N. (2007) Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science,* **316,** 1746-1748.

Mariotte, P., Meugnier, C., Johnson, D., Thebault, A., Spiegelberger, T. & Buttler, A. (2013) Arbuscular mycorrhizal fungi reduce the differences in competitiveness between dominant and subordinate plant species. *Mycorrhiza,* **23,** 267-277.

Mensah, J.A., Koch, A.M., Antunes, P.M., Kiers, E.T., Hart, M. & Bucking, H. (2015) High functional diversity within species of arbuscular mycorrhizal fungi is associated with differences in phosphate and nitrogen uptake and fungal phosphate metabolism. *Mycorrhiza,* **25,** 533-546.

Moran, P.A.P. (1948) Rank correlation and permutation distributions. *Proceedings of the Cambridge Philosophical Society,* **44,** 142-144.

Moran, P.A.P. (1950) Notes on continuous stochastic phenomena. *Biometrika*, **37,** 17-23.

Mulder, C.P.H., Bazeley-White, E., Dimitrakopoulos, P.G., Hector, A., Scherer-Lorenzen, M. & Schmid, B. (2004) Species evenness and productivity in experimental plant communities. *Oikos,* **107,** 50-63.

Munkvold, L., Kjoller, R., Vestberg, M., Rosendahl, S. & Jakobsen, I. (2004) High functional diversity within species of arbuscular mycorrhizal fungi. *New Phytologist,* **164,** 357-

364.

Oksanen, J., Simpson, G., Weedon, J., Solymos, P., Szocs, E., Legendre, P., Bolker, B., FitzJohn, R., Barbour, M., Bengtsson, H., Dunne, F.v. & Stier, A. (2016) vegan: CRAN relase 2.3-3. *Zenodo*.

Pagel, M. (1999) The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology,* **48,** 612-622.

Piélou, E., C (1975) *Ecological diversity*. John Wiley & Sons, New York.

Powell, J.R., Parrent, J.L., Hart, M.M., Klironomos, J.N., Rillig, M.C. & Maherali, H. (2009) Phylogenetic trait conservatism and the evolution of functional trade-offs in arbuscular mycorrhizal fungi. *Proceedings of the Royal Society B-Biological Sciences,* **276,** 4237-4245.

Pozo, M.J. & Azcon-Aguilar, C. (2007) Unraveling mycorrhiza-induced resistance. *Current Opinion in Plant Biology,* **10,** 393-398.

Sanders, I.R. & Rodriguez, A. (2016) Aligning molecular studies of mycorrhizal fungal diversity with ecologically important levels of diversity in ecosystems. *ISME Journal,* **10,** 2780-2786.

Savary R, Masclaux FG, Wyss T, Droh G, Cruz Corella J, Machado AP, Morton JB Sanders IR. Submitted**.** A population genomics approach shows widespread geographical distribution of cryptic genomic forms of the symbiotic fungus *Rhizophagus irregularis***.**

van der Heijden, M.G.A., Klironomos, J.N., Ursic, M., Moutoglis, P., Streitwolf-Engel, R., Boller, T., Wiemken, A. & Sanders, I.R. (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature,* **396,** 69-72.

van der Heijden, M.G.A., Martin, F.M., Selosse, M.A. & Sanders, I.R. (2015) Mycorrhizal ecology and evolution: The past, the present, and the future. *New Phytologist,* **205,** 1406-1423.

Vogelsang, K.M., Reynolds, H.L. & Bever, J.D. (2006) Mycorrhizal fungal identity and richness determine the diversity and productivity of a tallgrass prairie system. *New Phytologist,* **172,** 554-562.

Wagg, C., Jansa, J., Schmid, B. & van der Heijden, M.G.A. (2011a) Belowground biodiversity effects of plant symbionts support aboveground productivity. *Ecology Letters,* **14,** 1001-1009.

Wagg, C., Jansa, J., Stadler, M., Schmid, B. & van der Heijden, M.G.A. (2011b) Mycorrhizal

fungal identity and diversity relaxes plant-plant competition. *Ecology*, **92,** 1303-1313.

Wang, B. & Qiu, Y.L. (2006) Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza*, **16,** 299-363.

Wardle, D.A., Bardgett, R.D., Klironomos, J.N., Setala, H., van der Putten, W.H. & Wall, D.H. (2004) Ecological linkages between aboveground and belowground biota. *Science*, **304,** 1629-1633.

Wyss, T., Masclaux, F.G., Rosikiewicz, P., Pagni, M. & Sanders, I.R. (2016) Population genomics reveals that within-fungus polymorphism is common and maintained in populations of the mycorrhizal fungus *Rhizophagus irregularis*. *ISME Journal*, **10,** 2514-2526.

Supplementary information

Table S1 Coefficients of linear regressions between relative contribution of each plant species to total aboveground dry mass and plant community evenness.

Plant species	Coefficient	SE	P-value
A.elatius	24.65	6.41	<0.001
F.pratensis	-105.27	-19.77	<0.001
T.pratense	48.94	5.69	<0.001
L.corniculatus	15.19	2.38	<0.001
K.arvensis	11.25	1.49	<0.001
P.vulgaris	5.54	0.8	<0.001

Figure S1 Relationships between ADM responsiveness and AMF colonization rate for each of the six plant species. See main text for statistics

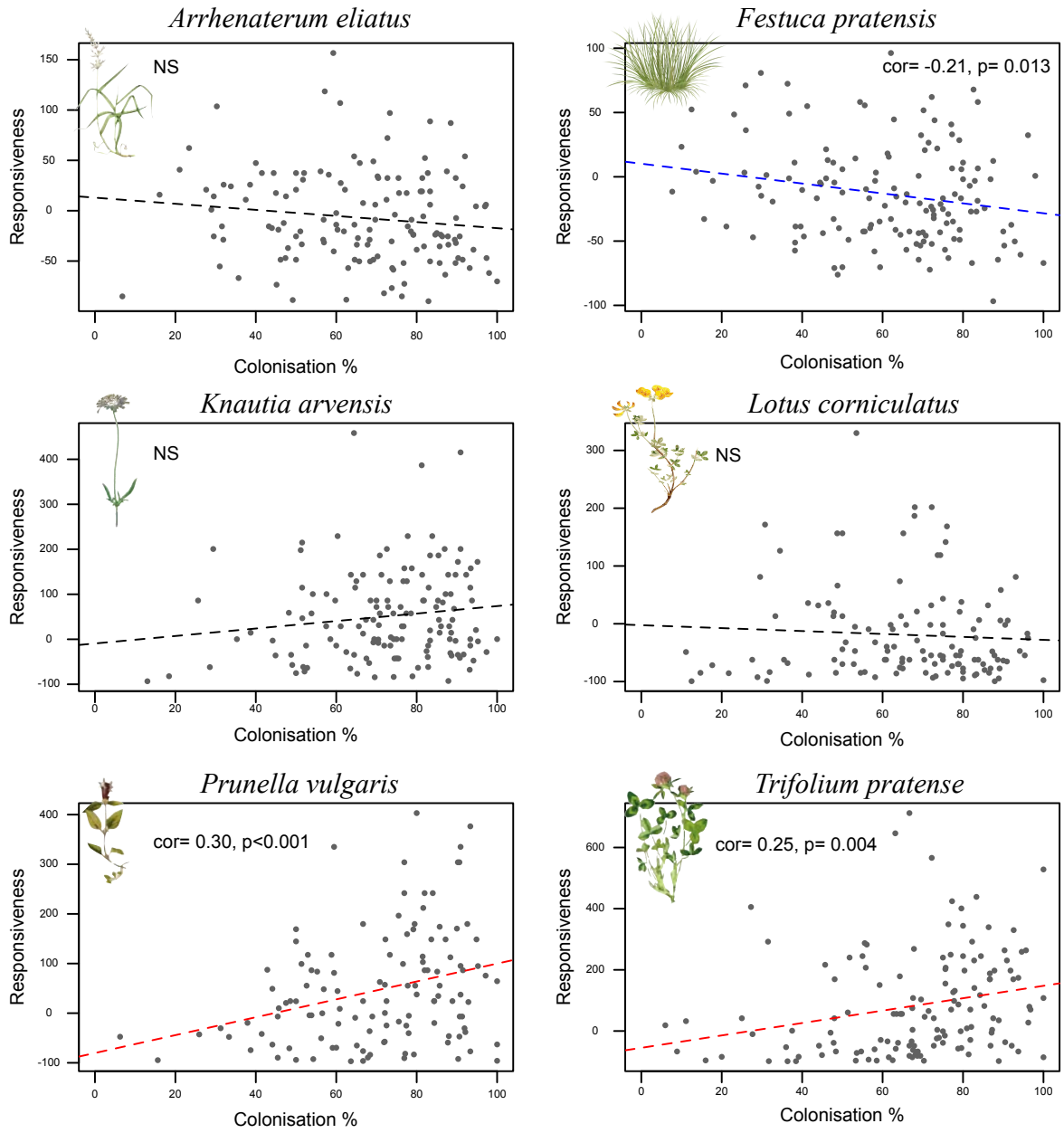


Figure S2 Plant community metric for all the treatments and NM for **(a)** ADM and RDM, **(b)** mesocosm productivity (TDM) and **(c)** plant community evenness.

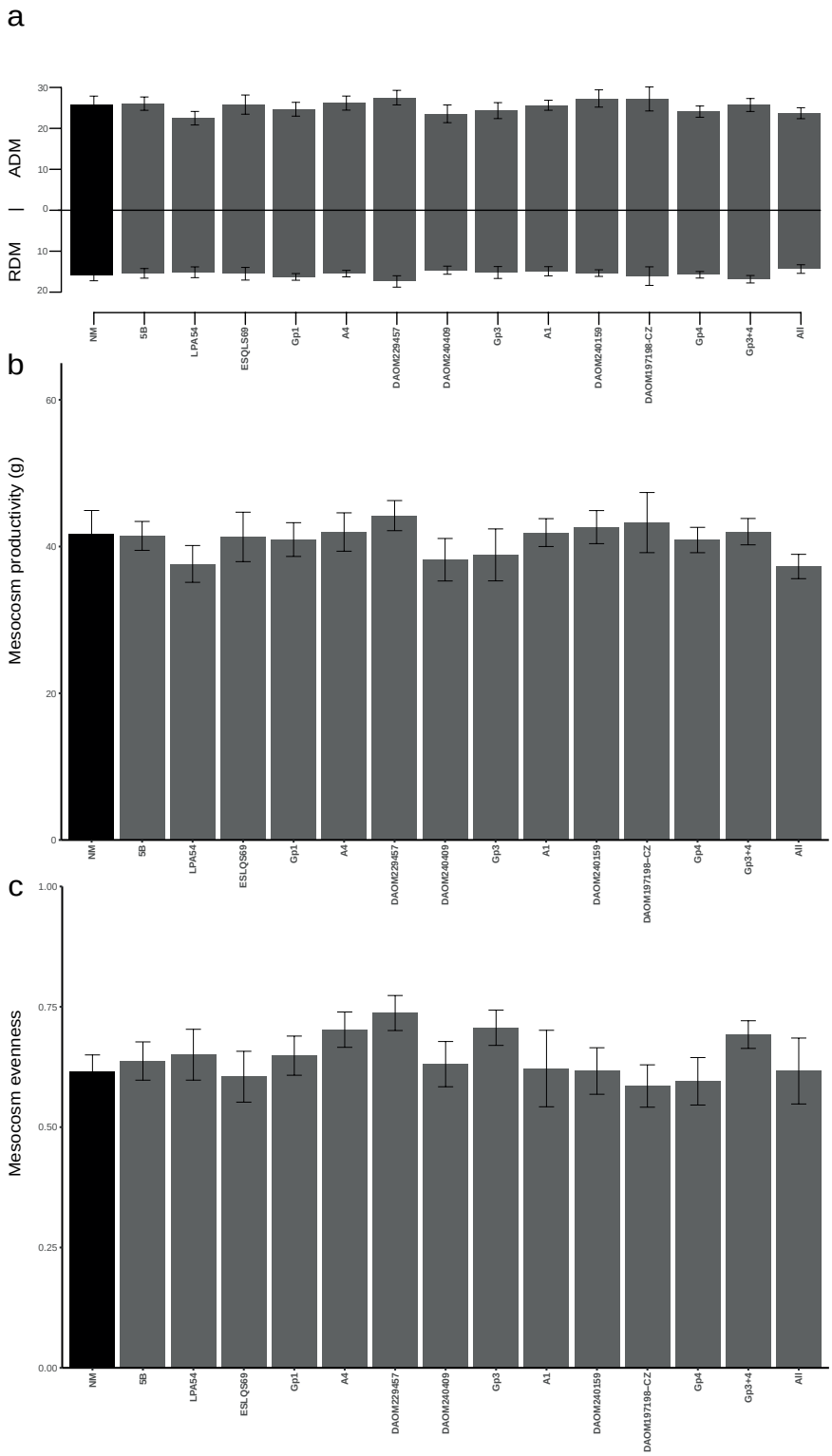


Figure S3 Relationship between plant community evenness and mean colonization rate for each mesocosm. See main text for statistics.

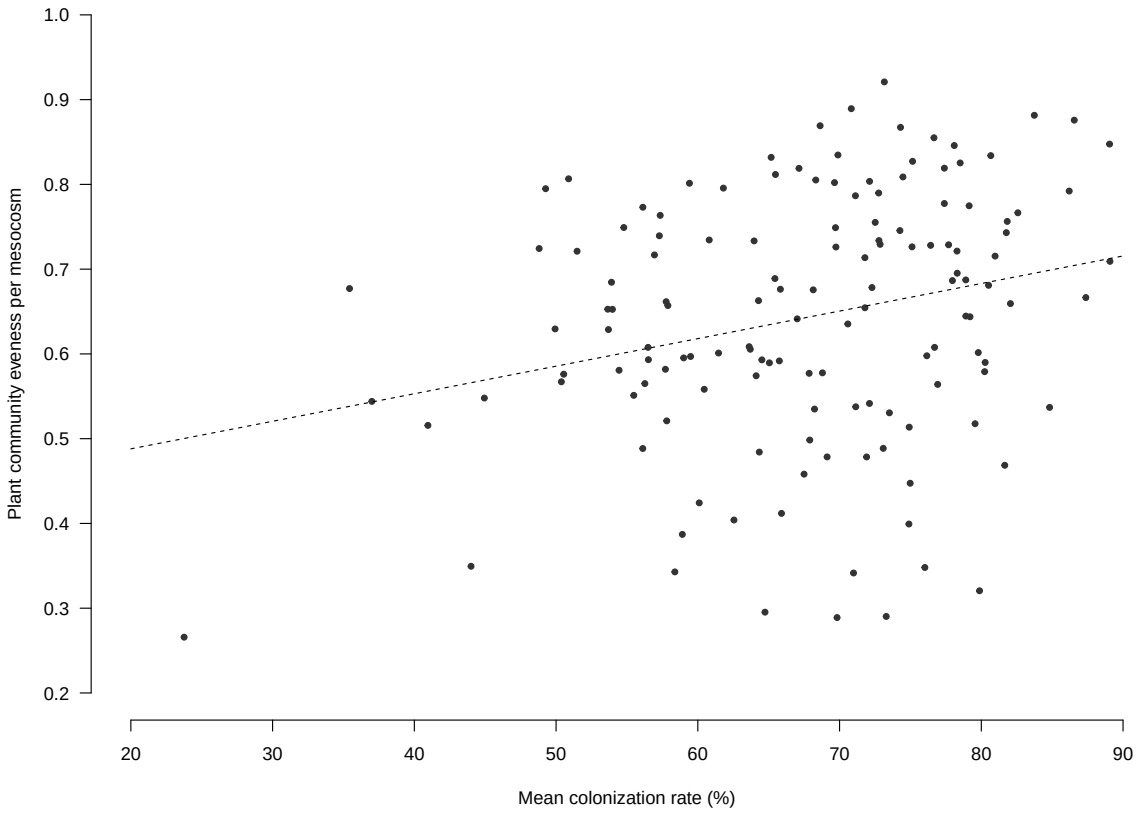
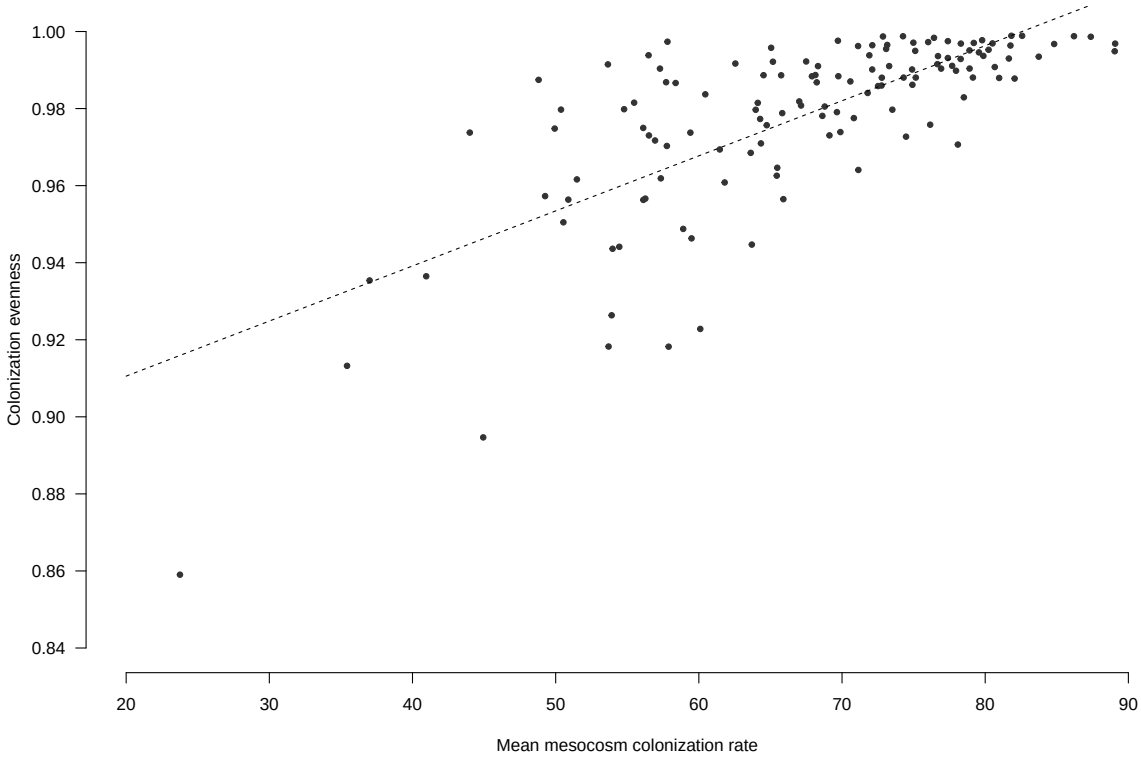


Figure S4 Relationship between colonization evenness and mean colonization rate for each mesocosm. See main text for statistics.



Supplementary Note S1: Roots straining and colonization measurement protocol

The staining protocol was as follow, the roots of all species were rinsed with clean water and transfer in a 2ml tube that was filled with 10% KOH and heated at 90°C for 1h for the roots of *T. pratense, A. elatius, F. pratensis* and *P. vulgaris*, 45min for *K. arvensis* and 30 min for *L. corniculatus*. During the heating time the KOH was change at least three times. At the end of the heating time, the KOH was removed and HCL 1% was added for 3 to 5 minutes. Trypan blue solution was then added either for overnight staining or for a 2h staining at 90°C. The trypan blue was then removed and replaced by lactic acid at 80% for permanent conservation. Roots of each of the 900 plants were randomly spread on a petri dish previously marked with 100 cells of 4x4mm arranged in a grid pattern. The presence or absence of fungal structure was recorded on each cell containing a root and sums up to calculate the colonization level (%)

Chapter 4

Effects of intra-specific variation in *Rhizophagus irregularis* on the transcriptomes of mycorrhizas and Cassava in symbiosis

Romain Savary[1], Cindy Dupuis[1], Frederic Masclaux[1, 2], Ivan Mateus Gonzalez[1, 3], Edward Camilo Rojas Tayo[1, 4] and Ian R. Sanders[1].

[1] Department of Ecology and Evolution, Biophore Building, University of Lausanne, 1015 Lausanne, Switzerland. [2] Vital-IT Group, Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. [3] Department of Plant and Environmental Sciences, Thorvaldsensvej 40, 1871 Frederiksberg C, University of Copenhagen, Denmark. [4].

Abstract

Arbuscular mycorrhizal fungi (AMF) are known to promote plant growth of important crop plants such as rice or cassava. However, the variability in plant response to different isolates of a single species was shown to be high.

In order to more deeply understand, on both plant and fungal side, the range of responses and molecular mechanisms involved in this agriculturally important symbiosis, we conducted a greenhouse experiment. We inoculated 12 *Rhizophagus irregularis* isolates from 4 distinct genetic groups on clonal plants of the Cassava (*Manihot esculenta* Crantz, variety NGA-16). Transcriptomic data of both organisms obtained by the RNA sequencing of the roots were analysed using differential expression analysis and co-expression network analysis.

A conserved global cassava root reprogramming was observed during the symbiosis with any genetic groups of *R. irregularis.* However, on the fungal side a strong transcriptomic difference was found among the four genetic groups. We detected high activity of the fatty acid biosynthesis and proteolysis in the plant that we further investigated. We also found a fungal secreted effector, a chitin deacetylase, that probably allows the fungus to hide from the plant. These results highlight a low specificity of the plants molecular response to each genetic group, and a high variation in transcriptome expression within a single AMF species.

Key words: AMF, Chitin deacetylase, Dual RNA-seq, Fatty acid biosynthesis, *Manihot esculenta* (cassava), *Rhizophagus irregularis*, symbiosis.

Introduction

Cassava (*Manihot esculenta* Crantz) is one of the most important crop plants for tropical countries, feeding more than 800 million people (FAO, 2013). With the constant and exponential increase of the world population there is a need to increase the yield of this crop. It has been shown that the production of *M. esculenta*, in symbiosis with *Rhizophagus irregularis* one fungal endosymbiont, could be a possible solution in order to increase its yield (Ceballos et *al.*, 2013). Arbuscular mycorrhizal fungi (AMF) have been known for decades to alter plant growth, including major crops plants (Angelard et *al.*, 2010). More precisely *R. irregularis* and other AMF, produce an extensive network of hyphae around the roots of three quarters of terrestrial plants (van der Heijden et *al.*, 2015) that help plants to acquire nutrients such as phosphate and nitrate (Karandashov & Bucher 2005, Govindarajulu et *al.*, 2005). Endocellular structures within the plant roots, called arbuscules, exchange these nutrients with the plant for photosynthetic carbohydrates such as sucrose (Parniske et *al.*, 2008). Other benefits that plants obtain from the AM symbiosis are, increased plant resistance to insects, nematodes, bacterial or fungal pathogens (for a review see Pozo and Azcon-Aguilar, 2007), increased drought resistance (Al-Karaki et *al.*, 2004) and increased salinity tolerance (Feng et *al.*, 2002, Hajiboland et *al.*, 2010).

Up to now, responses of *M. esculenta* to inoculation with AMF were mainly characterized by plant growth variables (Ceballos et *al.*, 2013). However, RNA-seq is a good way, not only to obtain a more detailed view of the gene transcription of the host plant but also a better understanding of the gene transcription on the fungal side of the symbiosis. By conducting RNAseq on RNA from the roots, the transcriptome of the fungus is also available to be analyzed in a so-called dual RNA-seq study.

Despite the ecological importance of the AMF symbiosis for the large majority of land-plants (van der Heijden et *al.*, 2015) and for potential agricultural application (Ceballos et *al.*, 2013, Rodriguez & Sanders 2015), very few studies evaluated both plant and AMF transcriptomic responses during symbiosis (Handa et *al.*, 2015, Shu et *al.*, 2016). To our knowledge, no study to date jointly assessed both transcriptomic responses of AMF and *M. esculenta*.
The large majority of recent plant-AMF RNA-seq studies focused on the plant response, e.g. differential reprogramming of rice root types (Gutjahr et al., 2015, Fiorilli et al., 2015), leaf and fruit development in tomato (Cervantes-Gámez et al., 2015, Zouari et al., 2014), or plant

response, with or without AMF, to abiotic stresses such as K+ deprivation (Garcia et al., 2017) or atrazine (Song et al., 2015). However, at least two studies have focused on *R. irregularis* (Tsuzuki et al., 2016, Sugimura and Saito, 2017), disentangling the molecular mechanism of the fungi during plant colonization and under high levels of phosphate.

It was recently shown that inoculation with a variety of AMF isolates of a single species such as *R. irregularis* resulted in a wide range of plant growth response in cassava (Ceballos et *al.,* in prep.), rice (Angelard et *al.,* 2010), as well as in other plant species (Koch et *al.,* 2006, Mensah et *al.,* 2015). These types of plant response effects due to AMF intraspecific variability were also observed with other species of AMF (Munkvold et *al.,* 2004, Mensah et *al.,* 2015, Koch et *al.,* 2017). In order to understand the molecular mechanism behind this variability, as well as its potential role and impact on gene transcription during plant symbiosis it is, thus, important to include the AMF intraspecific diversity while studying the response of crop plants to such microorganisms.

In this context, recent deep molecular characterization of *R. irregularis* intraspecific diversity (Savary et *al.,* submitted) uncovered four main genetic groups. This newly discovered diversity is, therefore, of interest for plant ecologists and for potential future agronomic applications.

We investigated the morphological and molecular responses of both *M. esculenta* and *R. irregularis* during their symbiotic phase, and evaluated the potential functional effects of genetic variation of *R. irregularis*. We performed a greenhouse experiment where we inoculated clonal cassava plantlets with a selection of 12 *R. irregularis* isolates, spread across the four different genetic groups described in the phylogeny of this species (Savary et *al.,* submitted). Dual RNA-seq was performed on clonal cassava roots to obtain whole transcriptome expression profiles of non-inoculated *M. esculenta* and *M. esculenta* inoculated with 12 different *R. irregularis* isolates. This dual RNA-seq was analysed by differential expression analysis as well as through the lens of co-expression network analysis.

With this design we tested the hypothesis that i) within AMF species diversity will affect differentially the transcriptome of the clonal cassava plants. This hypothesis was formulated in accord with the observed high variation in plant response to the within AMF species diversity. A second tested hypothesis was that ii) if the different genetic groups within an AMF species represented different ecological functional species we must found different

transcriptome expression. We then investigated what were these changes in cassava and *R. irregularis*.

Material & Methods

Fungal treatments

Twelve isolates, identified as belonging to the species *R. irregularis* (Savary et *al.*, submitted) were grown with Ri T-DNA transformed carrot roots in an *in vitro* culture system for a period of three and half month in order to produce enough spores for inoculation (St-Arnaud *et al.*, 1996). The selected isolates spanned the phylogeny of this species and represented the four different genetic groups described in *R. irregularis*. The isolate selected for the four groups were SAMP7, ESLQS69, LPA54, BEG140 and Israel (representing Gp1), BEG72 (representing GP2), C3, DAOM229457 and A2 (representing Gp3) and DAOM243181, DAOM240448 and DAOM197198-CZ (representing Gp4)(Fig. 1a). All isolates were maintained in identical *in vitro* conditions to avoid environmental effects. Details of isolate origin are found in Savary et *al.*, (submitted).

Plant material and experimental design

The *Manihot esculenta* cultivar NGA16 originated from Nigeria and was obtained from the International Center for Tropical Agriculture (CIAT; https://ciat.cgiar.org/). It was multiplied in *in vitro* conditions in order to produce 400 clones. After micropropagation (Santana et *al.*, 2009), the different clones were grown individually in glass tubes with an M1 phytagel solution (Santana et *al.*, 2009) with 14h of daylight (light intensity 100 µE.m-2s-1) at 25°C in a growth chamber. After one month, the non-contaminated plantlets were planted into 0.37L small pots in a steam-sterilized soil composed of Seedling substrate (Klasmann) and perlite (1:1) and then moved onto tables in the greenhouse with constant conditions (28°C, 70%RH and 16h daylight). Young plants were protected from full light with a mesh for five weeks of acclimation. The plants were then transferred to final 2L pots with a new sterilized soil composed of substrate S4 (Klasmann), perlite, quartz sand and clay (1:1:1:1) and placed on pallets in the greenhouse with the same conditions for one more month before inoculation. All soil used in this experiment was steam-sterilized at 105°C for three consecutive sessions of 20 min.

Among the remaining living clonal plantlets, 208 were randomly chosen and 16 plantlets were assigned for inoculation to one of the 12 fungal treatments or to the CTL treatment (Fig. 1b). Each of the twelve isolates was inoculated on a 3 month and one week old plant with 300

spores diluted in 10 ml of pure ddH_2O. One control non-mycorrhizal treatment (CTL) was established by inoculating 10 ml of pure ddH_2O without AMF. This inoculation process was repeated on sixteen clones of NGA-16 in order to obtain 16 replicates for each of the 12 fungal treatments and the non-mycorrhizal treatment (CTL). Two replicates per treatment were then randomly assigned to one of the 8 blocks and to one position within the block. Thus, each block contained all the treatments arranged randomly (Fig. 1b). The blocks were rotated every two weeks in order to avoid microclimate effects in the greenhouse. Once inoculated and placed in the block, the plants were not watered during three days in order to avoid the washing of the inoculum.

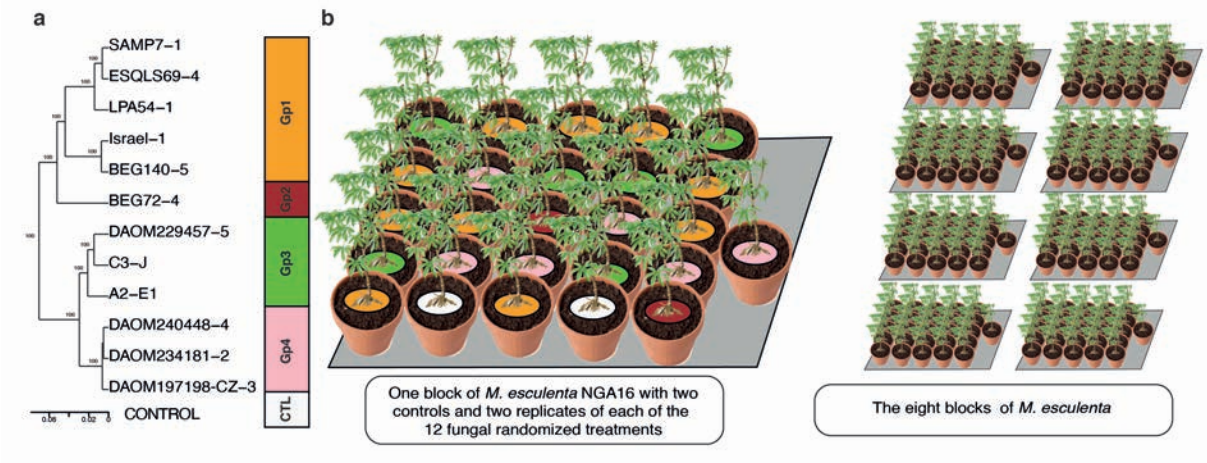


Fig. 1: ddRAD-seq phylogenetic diversity of *R. irregularis* and schematic representation of the greenhouse design. (a) ddRAD-seq phylogenetic tree of the four genetic groups based on 15 229 SNPs, 1085 insertion, 1455 deletions and 14 MNPs, support of the tree was calculated with 5000 bootstraps. Colours represent the belonging each of the four genetic groups according to in Savary et *al.*, submitted and are used as colour coding for theses groups along the paper (b) Schematic representation of one bloc out of eight, with 26 cassava plants, containing a randomization of two replicates of each of the 12 fungal + 1 CTL treatments. The reference for the images retrieved from Internet and modified are available in the supporting information.

Harvest and measurements

After 4 months of growth with the treatments the 7 months and 2 weeks old plants were harvested in one week, block per block. Each plant was unearthed and roots samples were collected for RNA extraction in less than one minute. Three samples of small non-tuberized roots, typically around 10 pieces of roots from different depth, were cleaned with ddH_20,

collected in an eppendorf tube and kept in liquid nitrogen before being transferred to a -80°C freezer for later extraction. Another sample of non-tuberized roots was kept at -20°C for estimation of intraradical fungal colonization (Supp note S1). The plants were then dried in a stove for 72h at 60°C and then the above ground dry mass (ADM), the total root dry mass (RDM) and the tuberized root dry mass (TDM) were weighed (Table S1a).

Statistical analysis of M. esculenta phenotypic measurements

The ADM, the RDM, the TDM and the colonization for each *M. esculenta* plant were analyzed for differences among treatment using one-way ANOVA and significant difference between each treatment was assessed with Tukey HSD test.

To measure the AMF phylogenetic signal in plant quantitative growth traits, a phylogenetic tree was built using the variation among the deepest sequenced ddRAD-sequencing replicate of each of the 12 isolates (Savary et *al.,* submitted). We measured scalar distances (Wyss et *al.,* 2016) using the variation founded in the 49 348 covered positions in the genome among all samples. This variation is based on 15 229 SNPs, 1085 insertions, 1455 deletions and 14 MNPs. This tree was used to calculate a phylogenetic signal, using the phylosignal package (Keck et *al.,* 2016), in *M. esculenta* clones growing with the different isolates. The phylosignal allow us to measure the link between the *R. irregularis* phylogeny and the colonization rate, the RDM, the TDM and the ADM for 5 types of phylogenetic signal indicators: The Moran's I index (Moran 1948, 1950), The Abouheif's Cmean index (Abouheif 1999), Blomberg's K and K* (Blomberg et *al.,* 2003) and finally the Pagel's λ (Pagel 1999).

RNA extraction, RNA-seq library preparation and sequencing

RNA was extracted from 40 roots samples (31 inoculated roots, 9 CTL roots without AMF, Table S1b) following the protocol of Das et *al.,* (2013). All solutions were prepared with DEPC water, mortar and pestles were soaked overnight in a 50% bleach and ddH_20 solution, rinsed with ethanol and ddH_20. One hundred mg of fresh roots were ground in a mortar in liquid nitrogen. The powder was transferred in 800 µl of Extraction buffer (100 mM Tris-HCl pH 8.0, 10mM EDTA pH 8.0, 100mM LiCl, 1% SDS, 200mM β-mercaptoethanol). Samples were homogenized and then incubated for 5 minutes at room temperature. Four hundred µl of cold Phenol:Chloroform:Isoamyl alcohol (25:24:1; PCI) was added. After mixing, samples

were centrifuged at 14'000 rpm at 4°C for 15 min. The supernatant was collected and an equal amount of PCI was added. After mixing, samples were centrifuged for 10 min at 14'000 rpm and 4°C. The supernatant was collected and 1 volume of isopropanol and 0.1 volume of Sodium acetate (1M, pH 5.2) were added. After mixing, the samples were incubated for 10 min at room temperature and then centrifuged for 10 min at 14'000 rpm and 4°C. The pellet was washed with 500 µl ethanol 70%. It was then dissolved in 30 µl RNase-free water. Samples were treated with Macherey-Nagel Dnase kit and cleaned with the Nucleospin RNA cleanup XS kit from Macherey-Nagel. Concentration and integrity of the RNA samples were assessed both with Nanodrop 2000 and Fragment Analyser™ from Advanced Analytical, using RQN integrity score. For the remaining seven samples (Table S1b), grinding was performed and RNA extraction was achieved using the Maxwell™ 16 robot (Promega) with the Maxwell® 16 LEV Plant RNA Kit following manufacturers instructions.

Libraries were constructed by polyA selection with the RNA extracted from the root samples stored at -80°C. Each library was prepared with the TruSeq Stranded mRNA Sample Prep Kit from © Illumina, following the manufacturers protocol and by round of 4 libraries at the time. For each treatment, we chose RNA samples replicates according to quality, selecting the best ones in term of RQN score (Table S2), absence of degradation and contamination. Concentration of the libraries was assessed with both Quantifluor from Promega and quality with Fragment analyzer from Agilent (Table S2). Libraries were sequenced using Illumina HiSeq Technology on Illumina Hiseq2000 platform. 47 libraries were generated and six samples were pooled and paired-end sequenced (2X 100-nt) in eight separate lanes (Table S2). Each lane contained at least one control treatment and biological replicates were spread as much as possible across lanes in order to avoid lane effects. Thus, for each fungal treatment, the 3 to 4 replicates were spread into two to three lanes.

Bioinformatic pipeline

An overview of the pipeline employed for the analysis of both *M. esculenta* and *R. irregularis* can be found in Fig. S1. The quality of paired-end reads of all 47 libraries was first checked with FastQC and then Illumina adapters were removed from each library. Low quality nucleotides or reads smaller than 40bp were also removed with Trimmomatic version: 0.33 (Bolger et *al.*, 2014). A 4-base wide sliding window was applied in order to cut sections of reads with an average quality lower than 15 Phred.

The clean libraries were pseudo-aligned with the Kallisto pseudoaligner version 0.42.4 (Bray et *al.,* 2016) to obtain estimate counts based on an index of transcripts from both *M. esculenta* v6.1 transcripts (Bredeson et *al.,* 2016) and on *R. irregularis* N6 predicted genes (Lin et al., 2014, Mateus et *al.,* in prep.). The *tximport* function (Soneson et al., 2016) was used to create a datatable of estimated counts obtained from Kallisto in the R environnement (www.CRAN.R-project.org; R Development Core Team 2008).

A second strategy was applied by mapping the reads with the 2pass aligner STAR version 2.5.1b (Dobin et al., 2013) on both the *M. esculenta* genome v6.1 (Bredeson et al., 2016) and the N6 single nucleus genome of *R. irregularis* (Lin et al., 2014). The raw counts were then obtained from the .bam file using the command *featureCounts* from the Rsubread package (Liao et *al.,* 2014) with the *M. esculenta* v6.1 annotation file and a new annotation file based on a new gene prediction of the N6 *R. irregularis* genome (Mateus et *al.,* in prep.).

The four counts tables (two obtained from Kallisto for *M. esculenta* and *R. irregularis* and two obtained from STAR and *featurecounts* for *M. esculenta* and *R. irregularis*) containing the counts of the 47 libraries were then analyzed in parallel (Fig. S1). Global visualization of the data for both *M. esculenta* and *R. irregularis* was obtained by normalizing the Kallisto estimated counts with the variance stabilization transformations applied through the *varianceStabilizingTransformation* function and then by using the plotPCA function of DESeq2 (Love et *al.,* 2014). The differential analysis was run with the DESeq2 package (Love et *al.,* 2014) for transcripts in *M. esculenta* and genes in *R. irregularis*. Transcript for *M. esculenta* and gene for *R. irregularis* differentially express (DE) were kept only if present in the DESeq results following both methods (Kallisto and Star-*featureCounts*) of mapping and read counting.

Differential expression of M. esculenta transcripts

M. esculenta count tables were analyzed with the DESeq function, which takes raw counts and includes a normalization for library size. First, a global comparison including controls against all AMF treatments was performed in order to obtain *M. esculenta* transcript DE during symbiosis with AMF. The transcripts found to be significantly DE with the comparison CTL vs AMF were considered as the core transcripts for the *M. esculenta–R. irregularis* symbiosis. These transcripts were kept for gene ontology (GO) enrichment analysis as well as for volcano plot visualization of the most highly differentially expressed transcripts.

Second, in order to obtain a more precise list of core *M. esculenta* transcripts express during symbiosis with *R. irregularis* and commonly differentially expressed in each genetic group of *R. irregularis* as well as to obtain specific *M. esculenta* DE transcript in response to inoculation of *R. irregularis* to each genetic group, we performed the following comparisons; i) CTL vs Gp1, ii) CTL vs Gp2, iii) CTL vs Gp3, iiii) CTL vs Gp4. The intersect of each of these four lists of DE transcript gave us the second list of core genes, but this time each of these transcript are significantly DE in each genetic group compared to non-inoculated plant. This approach was repeated after removing the seven libraries that were extracted with the other method of extraction with the Maxwell^{TM} robot. These lists were used for Venn diagram construction as well as for GO enrichment analysis.

In order to confirm the efficiency of the fungal treatments, we searched in the main results 11 homologues of plant genes (Gst1, Lec5, Scp1, Pt4, Vapyrin, Flot4, Ann2, Lec7, Glp1, Ram2, Ha1) known to be differently transcribed during AMF symbiosis (Hogekamp and Küster, 2013).

Differential expression of R. irregularis genes

R. irregularis counts tables were also analyzed with the DESeq function similarly to *M. esculenta* count tables. In these analyses, the CTL treatments were removed, as they do not contain *R. irregularis* transcriptomes. Comparisons were performed among the four genetic groups in order to achieve each pair-wise comparison. Thus, GP1 was compared to GP2, GP3 and GP4. GP2 was compared to GP3 and GP4. GP3 was compared to GP4. The six lists of DE genes were used for GO enrichment analysis.

Proteins prediction, annotations and GO enrichment analysis

Prediction of fungal protein coding genes was performed with the tool Augustus based on hidden Markov model. Augustus was trained based on a dataset of *Aspergillus* proteins. The dataset of *Aspegillus* proteins was generated by retrieving proteins reviewed by Swiss-prot from www.uniprot.org. Scipio and BLAT were used to align the Aspergillus proteins on the N6 genome and to define the gene structure. The structure of the genes was used to train and optimize Augustus following Augustus instructions (http://www.molecularevolution.org/molevolfiles/exercises/augustus/training.html).

We generated hints for Augustus predictions from DAOM197198 RNA-seq data using STAR aligner and cufflinks. Augustus predictions with hints were performed on the repeat-masked N6 genome using the trained parameters for the species. The same procedure was applied for the predictions of the cassava proteins except that Augustus was trained with plant proteins.
R. irregularis N6 proteins were blasted against the UniProtKB fungal protein database (www.uniprot.org); the blast hits were then used in Blast2GO with the standard parameter (Conesa et *al.*, 2005) in order to obtain GO terms. The same procedure was applied for the proteins of *M. esculenta*, except that the proteins were blasted against the UniProtKB "viridiplantae" database. The GOseq package (Young et *al.*, 2010) was then used to perform GO enrichment analysis in accounting for gene length bias. An FDR was applied to detect the enriched GO terms (Benjamini & Hochberg 1995).

Co-expression analysis

In complement to the differential analysis, we performed a co-expression analysis with the package weighted correlation network analysis (WGCNA, Langfelder & Horvath, 2008). Such an analysis has the advantage over the differential expression analysis, that it can detect genes that potentially did not change significantly in mean expression between two conditions but that are central in the co-regulation network by highly correlated expression with a large number of other genes.
For this analysis we used the counts obtained from Kallisto of both species, each time we removed control libraries as well as the low count genes with a minimum of 1 read per libraries and finally we removed low variation genes. The counts were then normalized using the varianceStabilizingTransformation function of DESeq2 as suggested by Langfelder and Horvath (FAQ, WGCNA website, 2014). Ultimately, library outliers were removed based on hierarchical clustering (Analysis 1: Q13, C10, C13, G7, Analysis 2: Q13, C10, C13, G7).
We then performed this analysis in two variants. First we implemented a multi-species co-expression analysis using the filtered and normalized 949 DE transcripts of *M. esculenta* during symbiosis that we concatenated to all the filtered and normalized genes of *R. irregularis*. This analysis was performed only on colonized plants and, thus, the control plants were removed. The modules were built with a minimum size of 30 genes and were composed of plant DE transcripts and fungal genes. They were then correlated to the mean of the three phenotypic measurements (ADM, RDM, TDM) and to the colonization of each sample as well as to the mean colonization for each isolate that we will refer as colonizer type. Modules

of genes and transcripts, significantly related to morphological traits, were conserved for network analysis in order to detect hub genes (module membership >0.9). Gene lists of significant modules were measured for enrichment in GO terms.

Secondly, the 949 DEG transcripts of inoculated *M. esculenta* inoculated were assembled in co-regulation modules with a minimum module size of 5. Each module was then summarized into an eigengene values. In parallel, the same process was applied to the 10'727 filtered and transformed fungal genes with a minimum module size of 100. The eigengene values for each plant module was then correlated with the eigengene value of each fungal module in order to perform a module-to-module association. The gene lists for highly significantly correlated modules of *M. esculenta* and *R. irregularis*, were then used for GO enrichment. Hub genes (module membership >0.9) in these modules were then inspected.

Effectors

It was hypothesised that fungal effector proteins play a major role in the symbiosis with plant, by controlling the immune system of the host. Evidence showed that these effectors are highly conserved across AMF species (Sędzielewska Toro & Brachmann, 2016). The nucleotide sequences of the 64 effectors found by Sędzielewska Toro & Brachmann (2016) in *Rhizophagus clarus* were retrieved from NCBI (KU305736 - KU305799) and homologues were searched in the Nu6 *Rhizophagus irregularis* genome by retaining the best hit after Blastn. The presence of effector homologues of *Rhizophagus irregularis* was searched within the DE fungal gene lists and list of co-expressed gene modules.

Results

Phenotypic differences in fungal and plant traits

Colonisation by AMF of the cassava roots was significantly different among the 12 genetically diverse isolates of the AMF species *R. irregularis* across all treatment excluding the control. Some isolate were low colonizers (BEG72: 21.7%±4.5 SE) while other were higer colonizers (DAOM197198: 72.9%%±5.9 SE). Moreover this colonization was not randomly distributed among isolates but was correlated with their phylogeny, as shown by four significant and one marginally significant phylogenetic signal indicators (Fig. 2). Inoculation of plants did not result in differences in either RDM, TDM or ADM, either between the controls and the different isolates or among isolates. We did not detect phylogenetic signal in any of these values (Fig. 2). Phenotypic information for all the 208 plants and for plants used for RNA-seq is available in Table S1a and S1b.

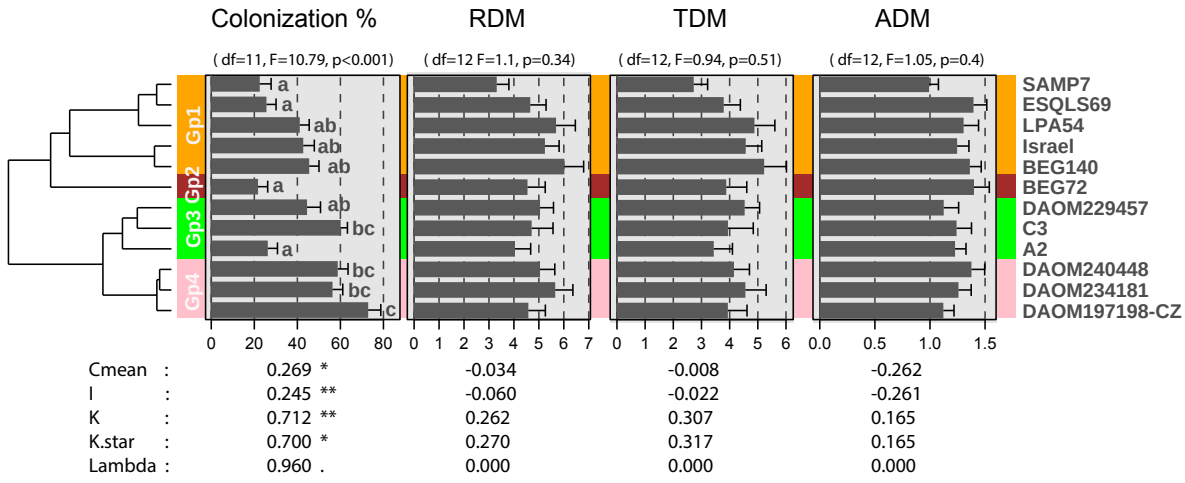


Figure 2: Phenotypic measurements of inoculated cassava plants in relation to phylogenetic relatedness of each *R. irregularis* isolate. Bars in the histogram represent the mean of each observed traits (Colonization level %, RDM, TDM and ADM) as well as the standard error are represented. One-way ANOVA results for each phenotypic trait are indicated above each bar. The One-way ANOVA on colonization did not included the controls. Conversely, controls were included in One-way ANOVA for the other traits (RDM, TDM, ADM). Significant differences were determined by post-hoc Tukey HSD and results are represented by letters above bars. Underneath each barplot, results for five phylosignal indicators calculated based on the mean of each trait is given with their respective *p*-value (. <0.1, * <0.05, ** <0.01).

RNA-seq

RNAseq generated a total of 4.47 billion reads with an average of 95 million of reads per library (Table S2). Information about the STAR 2pass mapping, counts measurement with *featureCounts* for both species could be found in Table S3, S4, S5 and S6. Mean pseudo-mapped reads with Kallisto to the *M. esculenta* genome was 8'304'761 ±1'353'166 (Table S7) and to the *R. irregularis* genome 1'135'415±169'133 SE reads (Table S8). After trimming and mapping with both methods, removing contaminated controls, wrongly assigning treatments (Fig. S2), low quality libraries (CTL8, Q11, G12, G4), filtering transcripts and genes with less than one reads, we obtained *M. esculenta* expression information for 37 982 transcripts (Kallisto) and for 27 633 transcripts (STAR-featureCounts) out of a total of 41 381 transcripts. *R. irregularis* expression resulted in information on 13 107 genes (Kallisto) and 12 599 (STAR-featureCounts) out of 15 953 genes. The raw reads of the 47 libraries are available on the ncbi bioproject SRAXXXXXX.

M. esculenta *response*

We observed a reprogramming of cassava roots whole transcriptome while in symbiosis with any *R. irregularis* isolate or genetic groups (Fig. 3a). However cassava plants did not respond globally differentially to the different *R. irregularis* genetic groups. We found 949 differentially transcribed genes between AMF treatments and the controls (845 up regulated, 104 down regulated). 73% of the transcripts were common to both types of mapping and read counting analysis (Kalisto DE transcripts: 1161, 2passSTAR-*featureCOunts* DE transcripts: 1089, Fig. 3b). We found 10 significantly differentially expressed genes out of the 11 common up regulated plant gene during AMF symbiosis (Fig. 3b and c).

To obtain a more precise view of the core gene plant toolkit used during the symbiosis with *R. irregularis*, we retrieved the 353 plant genes that are differentially transcribed with any isolate genetic group compare to the CTL. Finally, in order to avoid any bias towards the seven samples extracted with the robot, we removed those samples and retrieved again the DE transcripts across the four comparisons between CTL and four genetic groups, which resulted in 310 DE transcripts. The three lists of gene can be found in Table S9a, S9b and S9c, along with the DESeq2 results including their $\log_2$fold change and adjusted p-value for the Kallisto counts data.

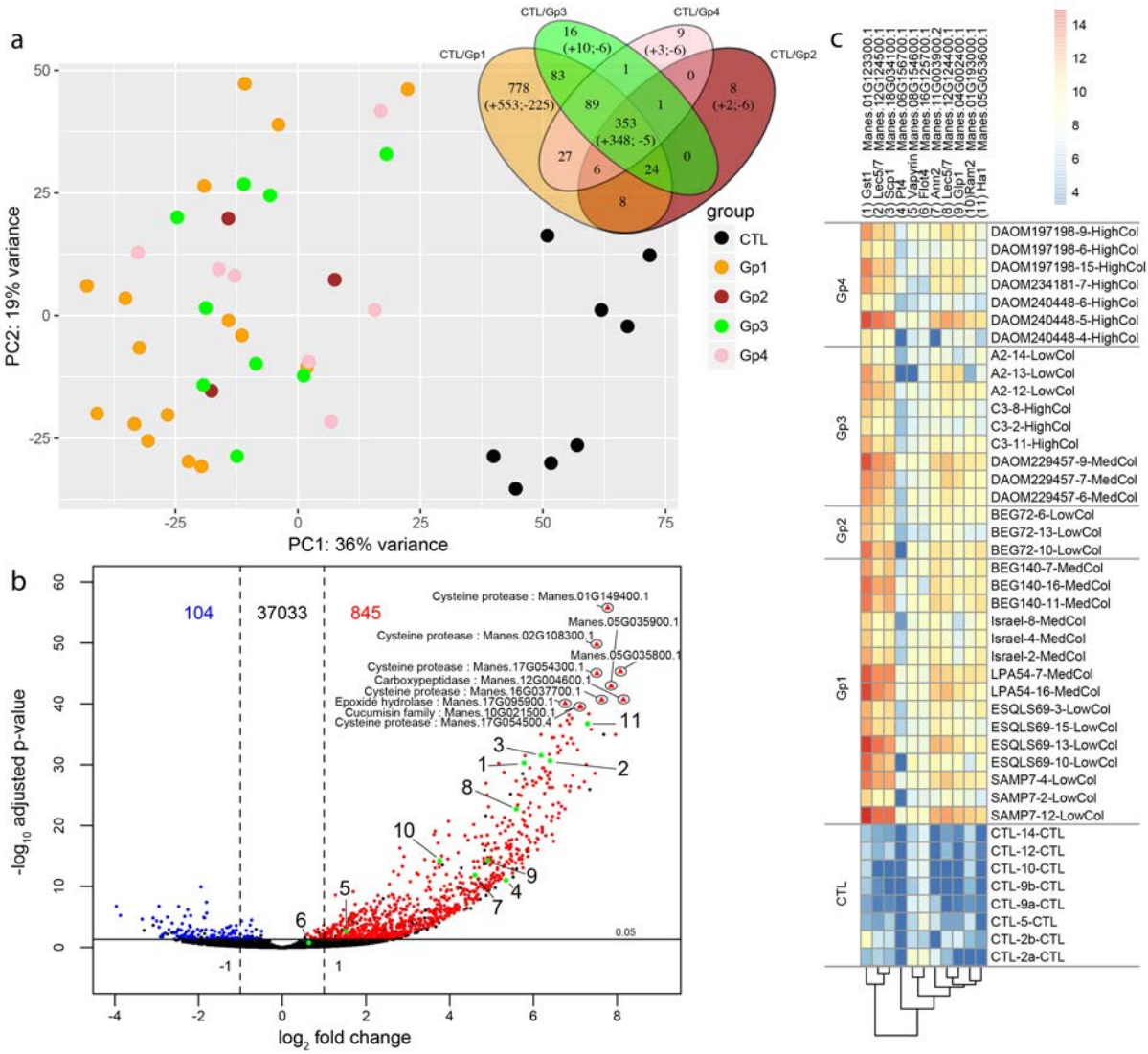
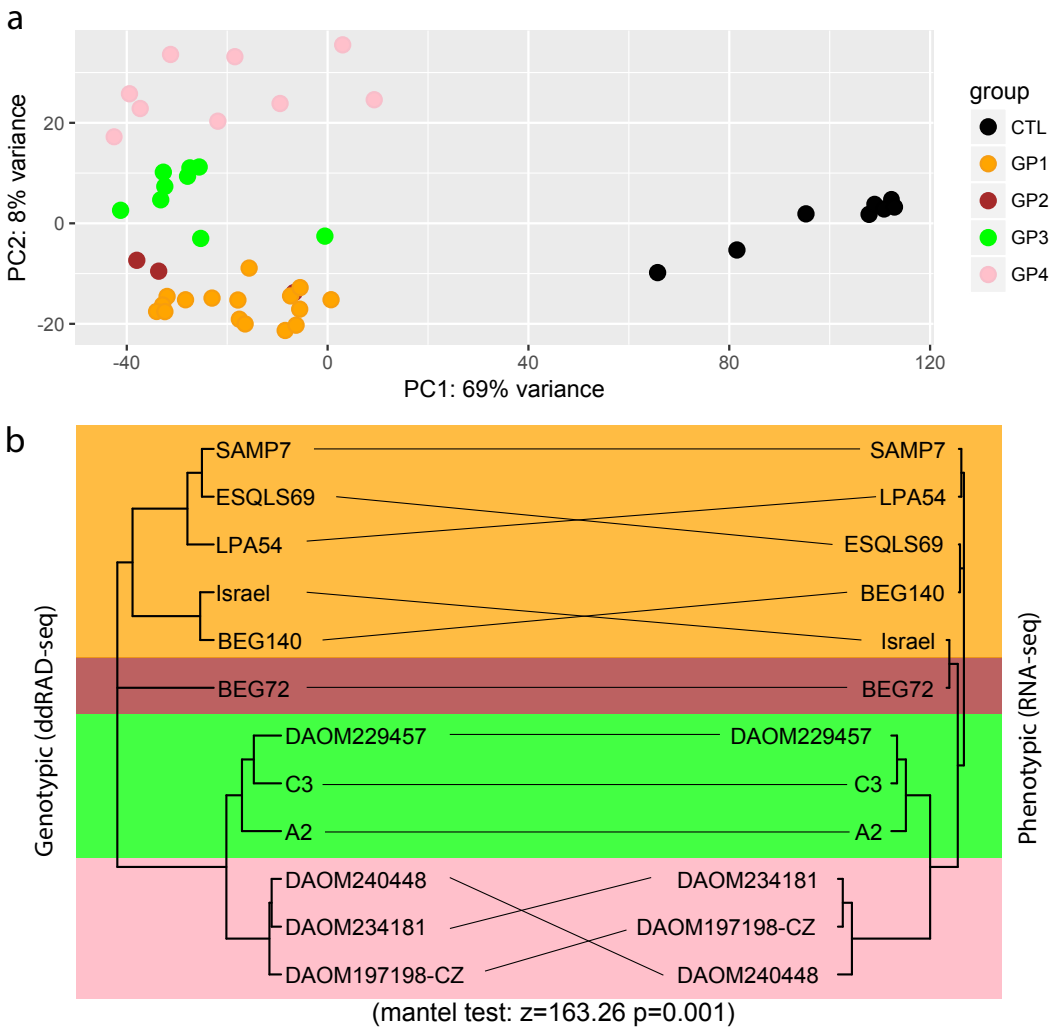


Figure 3: *M. esculenta* transcriptomic response to the inoculation with 12 *R. irregularis* isolates from 4 different genetic groups. (a) PCA based on the normalized (*vst*, DESeq2) counts expression (Kallisto) of 37 982 transcripts from control (n=8) and inoculated (n=35) plants. Venn diagram represent shared (353) or specific (778, 16, 9, 8) DE transcripts between controls and each genetic group. (b) Volcano plot of the comparison CTL vs AMF made with Kallisto counts, with 37 982 transcripts representing down-regulated (blue, left, 104) and up-regulated (red, right, 845) transcripts found in both type of mapping analysis (Kallisto-DESeq2 or 2passStar-featurCounts-DESeq2). Black dots (37 033) represent either non-differentially express transcript in both mapping strategies or differentially express transcript in one mapping strategy but not found in the other. The ten transcripts with the highest $-\log_{10}$ adjusted p-value are represented by circle triangle and annotated with their protein annotation and gene code. Green dots and numbers represent cassava homologues of eleven plant genes commonly express during AMF symbiosis. The name of each of these gene is found in the heatmap (c). (c) Heatmap detailing the normalized expression of the eleven cassava plant genes commonly DE in other mycorrhizal plants, for the 35 inoculated and 8 controls plants.

R. irregularis *transcriptiome diversity*

The PCA of *R. irregularis* indicated that the four differential genetic forms of *R. irregularis* have different expression pattern during the symbiosis with cassava (Fig. 4a). The differentiation of global expression among the 4 genetic groups along PC2, correlated strongly with their ddRAD-seq phylogeny (mantel test: z-value=163.26, p=0.001, Fig. 4b). The number of differentially expressed genes increased with increasing genetic distance (Fig. 4c). The results for the comparison between each genetic group as gene list and DESeq2 results for Kallisto counts can be found in Table S10a, b, c, d, e and f.

Genes differentially transcribed in only one fungal genetic groups compare to other were found (Fig. 4d) for each of the four group, with 27 genes differentially transcribed in GP1, 8 differentially transcribed in GP2, 12 s differentially transcribed in GP3 and 46 differentially transcribed in GP4.



c

	GP1		GP2		GP3		GP4
GP1	-		75 (+43;-32)		503 (+179;-324)		772 (+317;-455)
GP2	75 (+32;-43)		-		85 (+19;-66)		196 (+41;-155)
GP3	503 (+324;-179)		85 (+66;-19)		-		153 (+47;-106)
GP4	772 (+455;-317)		196 (+155;-41)		153 (+106;-47)		-

d

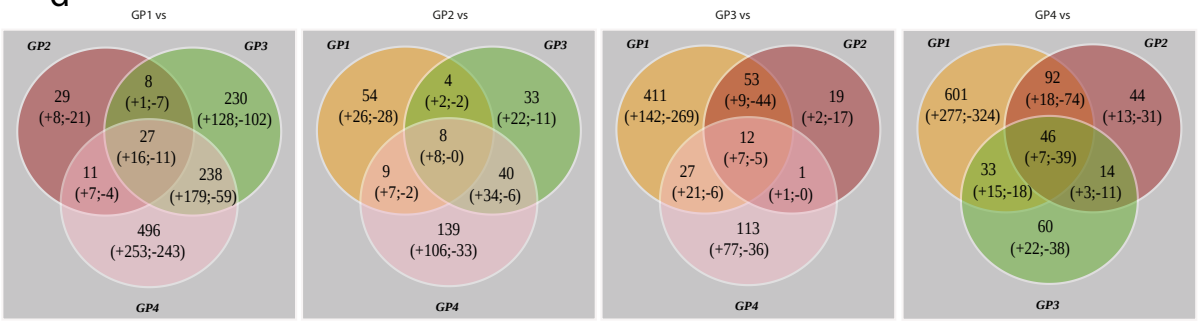


Figure 4: Transcriptomic response of 12 isolates of *R. irregularis* from 4 genetic groups to the inoculation on *M. esculenta*. (a) PCA based on normalized (*vst*, DESeq2) counts expression (Kallisto) of 13 107 genes from 37 *R. irregularis* transcriptomes from 12 isolates and their three replicates. The eight controls were included to show the low level of *R. irregularis* reads contamination. (b) Hierarchical clustering of the genetic distance and hierarchical clustering of the expression of 12 isolates of *R. irregularis*. The genetic hierarchical clustering was produced out of the scalar distances between the best replicates of 12 isolates of *R. irregularis* based on 15 229 SNPs, 1 085 insertions, 1 455 deletions and 14 MNPs obtained by ddRAD-seq. The expression hierarchical clustering was performed using as distance the mean of the PC2 value across replicates per isolate. The result of a mantel test applied on both distances matrices is shown (a). (c) Number of differentially expresses genes shared in both mapping and read counting strategies, across the comparisons between every *R. irregularis* genetic groups. Number in bracket indicated the number of up-regulated and down-regulated genes. (d) Venn diagram indicates the number of specific gene up and down-regulated in each genetic group.

Cassava and R. irregularis *GO enrichment*

Gene ontology terms enriched in the DE transcripts of the inoculated cassava plants and found in common in the three gene lists (949, 353, 310) were proteolysis, serine-type endopeptidase activity and the cysteine-type peptidase activity. Fatty acid biosynthetic process as well as ammonium transmembrane transport, carbohydrate binding and transferase activity were also enriched (Fig. 5a and b, Fig. S3, Table S11).

Among the different fungi, the GO enrichment of the 6 DE gene lists, resulted in a main GO term, the "oxidoreductase activity" for 3 comparisons between the most genetically distant fungi (GP1 vs GP3, GP1 vs GP4, Gp2 vs GP4, Fig. S4, Table S12a, b, c, d, e and f). All other comparisons resulted in no other enriched GO term, except for the most distant comparison between isolates from GP1 and GP4. In this DE gene list, we also detected enrichment for cell wall organization and biogenesis, cellular organization and carbohydrate metabolism processes (Fig. S4 and Fig. S5, Table S12a, b, c, d, e and f).

Multi-species co-expression networks and traits

The first WGCNA analysis of the concatenated gene expression of the plants and fungi resulted 35 modules with gene sets containing between 46 and 1811 genes (Fig. S6a, b, c and d). We found after FDR correction one module (lightcyan, n=214, Fig. S6b, Table S13a, b and c) that was significantly correlated with colonizer type (cor=0.73, uncorrected p=3*10^-6, corrected p=0.00054) and that same module showed the highest value among the other modules (cor=0.53, uncorrected p=0.002, corrected p=0.36) for correlation with other colonization measurements. Another module (purple, n=378, Fig. S6b, Table S14a, b and c) was found to be significantly correlated, after FDR correction, to the colonizer type (respectively: cor=0.67, corrected p=0.0072). The ligthcyan and purple modules exclusively contained fungal genes with ~40% and 52%, of characterized proteins, respectively. In the modules, we found 22 and 11 hub fungal genes in the network with a module membership value higher than 0.9 (or kME>0.9) respectively. Among these hub genes we found only in the "lightcyan" module, 8 genes that were significantly linked to the colonizer type. These genes are mainly uncharacterized proteins, except for Tos3p (g7882.t1) and ring-8 (g15266.t1, Table. S13) The GO term enrichment of the two modules gave significant term after p-value correction (FDR) only for the purple module, for oxidoreductase activity (padj= 5.5e-10). However the higher significant term in the lightcyan module without correction is also the oxidoreductase activity (p= 0.00779). The lightcyan module contained annotated genes that were differentially expressed between the most distant *R. irregularis* strains. Among those genes we found a sterol 14-demethylase, Tos3p a homologue of the effector Cdc15p, Cka2p, an alpha/beta hydrolase, ring8, a nitrogen reductase, a histone deacetylase, histone H2A, Env9, a Phosphatidylserine decarboxylase and a Glycerol-3-phosphate dehydrogenase

Co-expression of DE M. esculenta *transcript and* R. irregularis *genes*

In the second analysis with the WGCNA package, we obtained from the 949 *M. esculenta* DE transcripts, 5 modules containing from 9 (yellow module) to 720 (turquoise module) transcripts (Fig. S7 a, b and c). We obtained from the 10'727 fungal genes 21 modules containing from 157 (royalblue module) to 1559 (turquoise module) genes (Fig. S7 a, b and c). The blue plant module (n=80) was the module for which the eigengene values correlated the best with two fungal gene modules (brown; n=684, cor=0.83, padj=$1.05*10^{-6}$ and pink: n=479, cor=0.82, padj=$2.1*10^{-6}$). Two other plant-fungus modules correlated significantly, after FDR correction. First the blue plant module (n=80) correlated with the fungal turquoise module (n=1559, cor=0.59, padj= 0.042) and secondly, the yellow plant module (n=9) correlated with the fungal cyan module. The GO enrichment of the blue module did not highlight a particular function after FDR p-value adjustment (Table S15a). However in this module we found 10 hub genes (Fig. S7 d and e, Table S15b and c), such as eukaryotic translation initiation factor 3 subunit, insulin-degrading enzyme, acetyl-carboxylase alpha-CT, auxine response factor, ABC transporter family, transporters, Leucine-rich repeat kinase family isoform 1, phospholipid-transporting ATPase, kinase family peptidoglycan-binding domain-containing and universal stress. The GO enrichment of the brown and pink modules, resulted in the brown module of 11 terms mainly linked to a ribosomal activity for translation, other than that the term signal transduction, small molecule metabolic process, cytoskeleton as well as generation of precursor metabolites and energy were enriched (Table S16a). This last term was the only term enriched in the pink fungal module (Table S17a). We found 13 hub genes in the brown module and 3 in the pink module (Fig. S7d and e, Table S16b and S17b).

Fatty acid biosynthesis, Fat *genes, lipid transport and fungal lipase*

Fatty acid biosynthesis being the second most enriched GO term in genes DE between cassava control plants and cassava-inoculated plants (Fig. 5b, Table S11). We, thus, investigated more deeply the hypothesis that an the increase in synthesis of plant fatty acids (Bravo et al., 2017, Kamel et al., 2017), represents a source of carbon for the fungus. We, thus, investigated the expression of the differential key genes, of plant fatty acid biosynthesis, fatty acid elongation, transport and potential digestion in the fungus. The normalized expression of all key genes of the fatty acid metabolism, in relation to the colonizer type, and

to the fungal treatment of the cassava can be found in supplementary information in Fig. S8 and Fig. S9.

All the 12 genes up regulated in the GO enrichment of fatty acid biosynthesis are key genes involved in the activation and elongation of the plant fatty acid (Fig. 5a and b). Moreover, the FatM gene (Medtr1g109110), a palmitoyl-acyl carrier thioesterase shown to be only conserved in mycorrhizal plants (Bravo et *al.*, 2016) and responsible for C16:0 palmitic acid synthesis, was found with blastp as two homologues in cassava (Manes.06G063300.1 and Manes.14G109400.1, Table S18a). Both were differentially expressed when the plant was in symbiosis with any *R. irregularis* genetic groups (Fig. 5a and b). A third palmitoyl-acyl carrier thioesterase found in the annotation was never DE. The homologues of FatA and FatB proteins of *Ricinus communis*, the two common plant acyl-ACP thioesterases (Sánchez-García et *al.*, 2010), were searched with blastp in cassava (Table S18b). The homologue of FatA (Manes.13G049100.1) was mostly DE between CTL and GP1 (Fig. 5a and b). Moreover it was found in the plant yellow module with a high module membership of 0.95, which correlated to two fungal gene modules. In contrast, none of the three FatB cassava homologues were found to be DE or present in modules linked to fungal modules (Fig. 5a and b). The production of lysophosphatidic acid from the acyl-CoA is performed in the endoplasmic-reticulum (ER) by the glycerol phosphate acyl transferases (GPAT). One of these enzymes, RAM2, is conserved in mycorrhizal plants and is constantly differentialy expressed in cassava with all genetic forms of *R. irregularis*. Another ER GPAT (Manes.01G193000.1) was strongly up regulated in the presence of any *R. irregularis* isolates. Two other acytransferase were DE and could have a role in the next strep of fatty acid transformation, from lysophosphatidic acid to phosphoatidic acid.

We inspected several different lipid transporters potentially up regulated during the AM symbiosis and suggested as potential lipid transporters from the plant to the fungus (Kamel et *al.*, 2017, Bravo et *al.*, 2017). Four lipid non-specific transfer proteins were found to be conserved across 14 mycorrhizal plants, including *Medicago truncatula*, but not found in 9 other non-mycorrhizal plants (Delaux et al., 2014). Homologues of these lipid transfer proteins were searched with bastp and gene(s) with the highest hits in the cassava genome were conserved. However, none of either the 4 conserved or 28 other non-specific annotated lipid transfer proteins were found to be differentially expressed or co-expressed in identified plant modules. We then applied the same blasp procedure to an ABC transporter conserved across all the mycorrhizal plant species (Delaux et al., 2014), indeed ABC transporters were often found to be involved in lipid transport (Li et al., 2016). This ABC transporter

(Manes.11G015600.1, Delaux et al., 2014, Fig. 5a and b, Table S19a) was found to be highly diff express in the presence of any fungal genetic groups.

Six other ABC transporters were found to be differentially expressed with one comparison or the other. One of these transcript (Manes.11G063600.1) was found to be DE in three comparisons out of four (CTL/GP1, CTL/GP2, CTL/GP3, Fig. 5a and b). Among the 6 DE ABC transporters, Manes.01G234500.1 and Manes.05G060300.1 are hub (MM >0.9) or nearly hub genes (MM= 0.92 and 0.89) in the blue co-expression plant module (Fig. 5a and b). In the hub genes of the blue plant module, we are found another DE lipid transporter (MM>0.9), a Phospholipid-transporting ATPase (Manes.09G123600.1, Fig. 5a and b) DE with any *R. irregularis* genetic groups and correlated to different fungal modules, as well as with the colonization.

A pair of ABC-G (STR/STR2) transporters were shown to be indispensable for arbuscule formation, conserved across mycorrhizal plant and suggested as potential lipid transporters (Bravo et al., 2017). The homologues of the *Medicago truncatula* ABC transporter STR (Manes.06G004300.2) and STR2 (Manes.13G104000.1, Zhang et al., 2010, Gutjahr et al., 2012) were found in cassava (Table S19b). Only one of the two, the STR2 was found to be highly differentially expressed with any CTL vs genetic group comparisons (log_2 foldchange range: 4.9-6.2, Fig. 5a and b). The first one, STR was DE with any genetic group with the Kallisto pseudo-mapping method, but did not appear with the other mapping strategy. However none were found in any identified modules. Moreover, STR2 is the only annotated ATP-binding cassette transporter among 45 that is differentially expressed.

On the fungal side, in the GO enriched terms between GP1 and GP4 we found two genes involved in phospholipid transport (g2588.t1 and g5809.t1), both being up regulated in GP1 compared to GP4 and the second one being found in the pink module (MM= 0.77). Both fungal phospholipid-transport ATPases were well correlated with the plant phospholipid-transport ATPase (Fig. S10), more particularly the fungal phospholipid-transport ATPase of the pink module, that correlated well (cor=0.82, p=1.09*10^-9) with the *M. esculenta* hub phospholipid-transport ATPase of the blue module.

Among the 42 lipases found in the annotation of *Rhizophagus irregularis*, one lipase “Phospholipase D nuclease” (g9241.t1) was found to be highly differentially expressed between fungi from GP1 and GP4, as well as present in the lightcyan module, with a high modularity membership and a high and significant gene-trait correlating significantly with colonization (Fig. 5b and d). In the brown fungal module we found a (g8617.t1)

lysophospholipase, the third highest express lipase in mean expression (Fig. 5b, Fig. S11) and in the pink fungal another (g7289.t1) lysophospholipase, the fourth lowest expressed lipase (Fig. 5b, Fig. S11), both modules were highly correlating to the blue plant module of 80 differentially expressed plant transcripts.

The genes linked to fatty acid metabolism, and found in the ligthcyan module, are genes of the glycerophospholipid metabolism of fungi, such as the glycerol-3-phosphate dehydrogenase (g4087.t1, MM=0.70, gene-trait p-value=0.04). This gene was also DE between GP1 and GP4. The diacylglycerol cholinephosphotransferase (g8706.t1 MM=0.53, gene-trait p-value=0.003) and the hub gene phosphatidylserine decarboxylase 2 (g74.t1, MM= 0.93, gene-trait p-value= 0.118) were also DE between GP1 and GP4. In the gene set of the enriched GO carbohydrate metabolic processes, we found another DE gene of the glycerophospholipid metabolism encoding a glycerol kinase (g1094.t1), an enzyme using the same substrat as the glycerol-3-phosphate dehydrogenase and being negatively correlated with it (Fig. S12). Phospholipase D and diacylglycerol cholinephosphotransferase use the same substrate, the phosphatidyl-choline, the expression of both enzymes being positively correlated (Fig. S13).

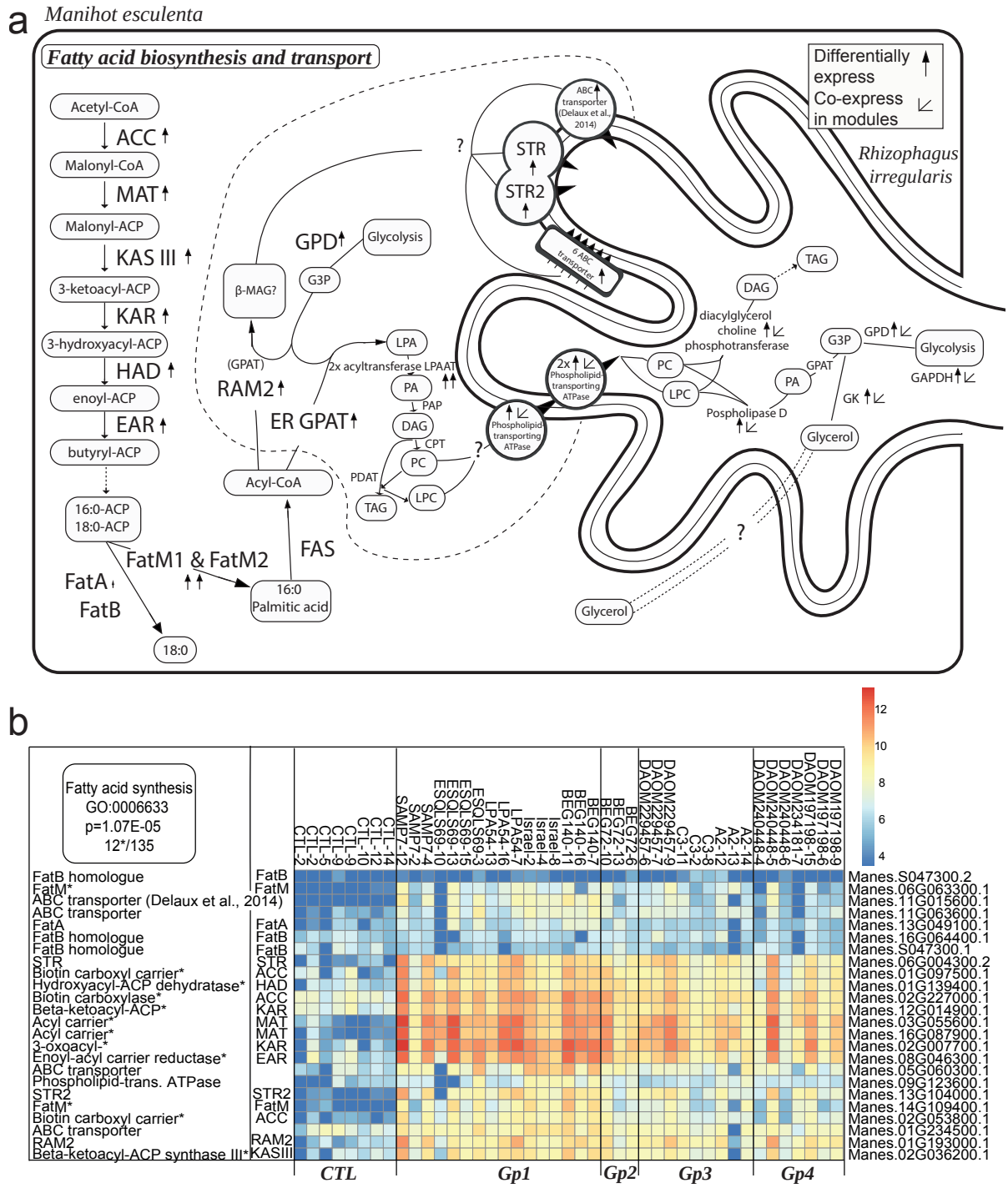


Figure 5: The fatty acid enriched pathway as a graphical representation, as a heatmap across the isolates and the control. (a) The fatty acid pathway from the formation and elongation in the plant to the suggestion of transport and use in the fungi. The acronym of the enzyme are, acetyl carboxylase (ACC), acyl-carrier-protein (MAT), Beta-ketoacyl-ACP synthase III (KASIII), 3-oxoacyl-(acyl-carrier-protein) reductase (KAR), hydroxyacyl-ACP dehydratase (HAD), enoyl-acyl carrier reductase (EAR), *Palmitoyl-acyl carrier protein* thioesterase (FATM1 and 2) , acyl-ACP thioesterases (FATA), acyl-ACP thioesterases (FATB), reduced arbuscular mycorrhization (RAM2), glycerol-3-phosphate acyltransferase (GPAT), stunted arbuscule ABC transporter (STR/STR2),

lysophophatidic acid (LPA), lysophosphatidic acid acyltransferase (LPAAT), phophatidic acid (PA), phosphatidate phosphatase (PAP), diacylglycerol (DAG), choline phosphotransferase (CPT), phosphatidylcholine (PC), phospholipid:diacylglycerol acyltransferase (PDTA), triacylglycerol (TAG), lysophosphatidylcholine (LPC), glycerol-3-phosphate (G3P), glycerol-3-phosphate dehydrogenase (GPD), glycerol kinase (GK), Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), chitin deacytelase (CDA2). (b) Heatmap of normalized transcript expression (Kallisto) for the 12 enriched genes of the fatty acid synthesis (*) as well as other genes in the fatty acid pathway.

Effectors

Among the 64 effector candidates conserved in different species of AMF (Sędzielewska Toro and Brachmann, 2016), 26 were found with a high blastn hit in the NU6 genome. Across all the genetic group comparison for DE genes, only the comparisons between the most divergent *R. irregularis* genetic groups (GP1 and GP4) revealed the presence of 2 effectors that were significantly differentially expressed (Table S20a). The same two effectors were found to be co-expressed with a high modularity membership in the "ligthcyan" (Table S20b) the best of the two fungal gene modules that correlated with colonizer type. The first most highly DE effector (g7109.t1, $\log_2$foldchange=2.94±0.88, padj=0.013), a chitin deacetylase (CDA2) was the only homologue (% identity 83.77%, e-value=7E-156, Fig. S9, Table S20c and d) in *R. irregularis* of the effector found in *R. clarus* (KU305765, Sędzielewska Toro and Brachmann 2016 supporting information, Table S20e). The other secreted effector differentially express is annotated as a Cdc15p (g2725.t1), a potential serine/threonine protein kinases.

Two high hit homologues of the well-described SP7 effector were found in the Nu6 genome (Table S21, g14902.t1 and g11683.t1). The first homologue (g14902.t1) had an identity of 100% to the first 64 amino acids of SP7. This transcript was thus considered as the homologue of SP7. This SP7 is DE between the two genetic groups GP1 and GP4 being up regulated in GP4 and down regulated in GP1 (log2 foldchange=3.69, Fig. S9). However, it was not found in any identified fungal modules. The ethylene response factor of *Medicago truncatula*, ERF19 shown to be the pathogenesis-related transcription factor with which the SP7 effector interacts in the nucleus of the plant was found in cassava. This homologue was not differentially expressed when plants were in symbiosis with AMF. Among another 35 ERF annotated in cassava, 6 were DE, and 2 (Manes.05G040000.1 and Manes.01G262100.1) were always DE with any genetic group. However, none were found in identified co-expression modules.

Discussion

In this study we show that despite no apparent differences in cassava growth in response to AMF, we can detect strong reprogramming of the plant root transcriptome. Such reprogramming was conserved across plants inoculated with different *R. irregularis* genetic groups, mainly in terms of the up-regulation of transcription in proteolysis genes, increased expression of numerous genes involved in the fatty acid biosynthesis, in ammonium transmembrane transport and in carbohydrate binding. We can, thus, reject the first hypothesis of an induction of major differential transcriptional changes due to within AMF species diversity. We proposed that changes in the plant phenotypes due to within species AMF diversity are more probably due to changes in a low number of targeted genes. Moreover we found that the colonization rate of the cassava roots with each isolate is tightly linked to the fungal phylogenetic relationship. This suggested that the fungus evolutionary history is impacting its ability to form the symbiosis. These findings suggested also that not only the plant control the fungus but also the fungus can control the plant though it evolutionary background.

Major transcriptomic differences were found between isolates and genetic groups of *R. irregularis*. These differences were clearly explained by the phylogenetic position of the isolate and the genetic groups, with up to around 5% of the genes differentially express between the two phylogenetically most distant genetic groups, GP1 and GP4. These findings tended to confirm the second hypothesis of possibly different functional species within *R. irregularis*. The transcriptomes differences among fungi from different genetic groups were mainly in expression of genes related to the oxidoreductase activity. Differences can be found as well between the two most phylogenetically distinct clades of *R. irregularis*, GP1 and GP4 in three other enrich GO terms, cell wall organization or biogenesis, cellular component and carbohydrate metabolic process. These change reflecting the major transitions of the transcriptome evolution between these distant clades of *R. irregularis*.

In the following sections we discussed the different major changes in cassava and in the different genetic groups of *R. irregularis*.

Insights in the chemical war for the control of the symbiosis

The first major and conserved change in cassava during AMF symbiosis was the up-regulation and enrichment of numerous genes involved in proteolysis. This gene arsenal possibly reflected the chemical war between the fungus attempting to control the plant host

immune system and the plant trying to avoid being controlled by the fungus. Indeed, in the plant, among the 10 most highly differentially expressed genes, five were cysteine proteases, the five other were, hydrolases, peptidases or uncharacterized proteins. It is known that one of the two classes of fungal effectors; the apoplastic effectors are cysteine-rich proteins (SCRs) such as avirulence factors avr2, avr4 or avr9 or FonSIX6 (Van der Hoorn et *al.*, 2001, Rooney et *al.*, 2005, Nui et *al.*, 2016). These effectors are secreted to reduce plant defences in the intracellular space (Sędzielewska Toro & Brachmann, 2016). Thus, these plant cysteine proteases were probably highly expressed to digest this kind of fungal effector. Similarly, plant carboxypeptidase, up-regulated during the symbiosis with *R. irregularis*, was probably involved in the fungal recognition as well as in the initation of the defence responses of the plant (Lui et al., 2008)

Among the 26 homologues of fungal effectors of *R. clarus* found in *R. irregularis*, two were found to be DE between the mostly distant fungal genetic groups and found to be co-expressed with 214 other genes in the lightcyan module that correlated with colonizer type. The first one, a chitin deacetylase (CDA2) could be central in the ability of the fungi to colonize strongly or not the roots of the cassava and potentially other plants. We suggested that this *R. irregularis* gene plays a role in hiding the presence of the fungus and its chitin components from the plant immune system by modifying the chitin into chitosan oligomers. This statement is strongly supported by evidence that the endophytic fungus *Pestalotiopsis sp.* inactivates the rice immune system by the modification of its chitin oligomers through a chitin deacetylase (Cord-Landwehr et al., 2016). One well-known system allowing the biotrophic symbiosis between plants and *R. irregularis* is the interaction between the SP7 fungal effector and the plant ethylene response factor (ERF19). SP7 inactivate plant defence related genes in the plant nucleus. As expected, the homologue of ERF19 was not DE. However we found that SP7 was only really active in the genetic group of the well studied fungus DAOM197198, the isolate used in the study of Kloppholz et al., 2011. SP7 was partially active in GP3 but not expressed at all in GP1 and GP2. Moreover, a higher expression of SP7 was associated with a higher colonization level (Kloppholz et *al.*, 2011). We also found that AMF colonization co-variated positively with the natural variation in the expression of SP7 in different isolate.

Recognition

The fourth most enriched plant term was carbohydrate binding. This term is composed of mainly kinases, such term reflect the ability of the plant to detect the fungi. Indeed, kinases

are often characterized as plant receptors of pathogens by detecting chitin, peptidoglycan or pathogen effectors. Among them, we found an L-type lectin-domain containing a receptor kinase, such kinase are known to be induced in response to fungal pathogens (Nivedita et *al.*, 2017) and were shown to play a role in plant pathogen resistance (Wang et *al.,* 2014). The homologues of the two common lectins 5 and 7 regularly reported as induced and specific to the AM symbiosis form part of this group of genes. Lectins are though to be able to bind fungal surface carbohydrates (Frenzel et *al.,* 2005). The other kinases detected probably fungal effectors or fungal chitin oligomers such as the LysM receptor-like kinase (Wan et *al.*, 2008, Antolın-Llovera et al., 2014). However the homologue of the *Ricinus communis* chitin elicitor receptor kinase (CERK1) was not found to be up regulated. Two other central potential cassava fungal receptors were found in the blue module. The first one is a leucine-rich repeat kinase family isoform 1 (LRR), similar LRR were shown to be essential for the activation of a non specific signal transduction in response to different microbes, the signal will then activate other symbiosis genes (e.g. SYMRK, Stracke et *al.*, 2002). The second one, a kinase family peptidoglycan-binding domain-containing, has the same functionality as peptidoglycan binding LysM gene.

The fungal partner also has to recognize the plant. We found an alpha/beta hydrolase linked to colonizer type varying in expression between genetic groups of *R. irregularis*. Alpha/beta hydrolases were shown in plant to be potential plant receptors of strigolactones (Gaiji et al., 2012), some being essential in plant such as DWARF14-like for initiating the AMF symbiosis (Gutjahr et al., 2015). Strigolactones are plant hormones and the detection of these hormones by the fungi is a requirement for efficient root colonization, we can then speculate that the fungus uses the same mechanisms as the plant for the recognition of strigolactones and that this fungal alpha/beta hydrolase is a strigolactone receptor.

Plant fatty acid synthesis, transport and recycling by fungi

The second most enriched gene ontology term is related to the increase of the regulation of numerous plant genes used for the synthesis of fatty acids. It was hypothesized quite early that AMF relied on their plant host for the production fatty acid, particularly for palmitic acid (Trépanier et al., 2005). We, thus, took advantage of expression data of the fungi and the variability across isolates to investigated more deeply the expression of genes potentially involved in this important pathway.

The common view of the plant-AMF symbiosis reported frequently this relation as being centred on the exchange of photosynthetic carbohydrates, mainly sugar in exchange for

nutrients such as nitrogen or phosphorous. The carbohydrates are thought to be used by the fungi to synthesize fatty acids (Trépanier et *al.,* 2005). However, as highlighted by Kamel et *al.,* 2017 and others, *Rhizophagus irregularis* as well as numerous other AMF species lack fatty acid synthases (FAS, Wewer et al., 2014, Tang et al., 2016), thus making the process of fatty acid synthesis impossible. It was also observed several times that plant under AMF symbiosis up regulate several lipids metabolism genes (Gomez et al., 2009, Tisserant et al., 2012). Moreover on the plant side, a specific and central gene of the fatty acid synthesis, the *Fat* gene (acyl-ACP thioesterase) *FatM* was described to occur only in mycorrhizal plants and was shown to be essential for the formation of healthy arbuscules (Bravo et al., 2016). *FatM* in combination with three other mycorrhizal plant conserved genes; *RAM2* and the ABC transporters *STR/STR2* were suggested to be the essential modules for the final synthesis of potential 16:0 β-monoacylglycerol (β-MAG) that are potentially then transported in the fungus (Bravo et *al.,* 2017). Another line of evidence is the change in lipid content when the plant enters into symbiosis (Wever et *al.,* 2014). Kamel et *al.,* 2017, ask the question of the origin of the palmitic acid found in the fungi and propose that the increased number of lipase genes in AMF, results from the requirement of the fungi to obtain carbon by digesting fatty acid transferred by the host plant.

In this study, twelve annotated genes of fatty acid synthesis were significantly up regulated between non-inoculated and inoculated cassava and with an expression pattern conserved across the inoculation of a wide genetic variety of *R. irregularis* isolates. The third gene, the most central in the blue plant module is an Acetyl-carboxylase alpha-CT (ACC). This enzyme was globably differentially expressed, but more importantly it is the third most central gene in the co-expression network of the DE transcript of the plant, linked strongly to two fungal gene modules (brown and pink) containing 685 and 479 fungal genes, respectively. This enzyme is central in the production of fatty acid as it catalyses the permanent carboxylation of actely-CoA into malonyl-CoA, the first product of the fatty acid synthesis. The other genes encoding for the enzymes that transformed actely-CoA into malonyl-CoA, the biotin carboxyl carrier and the biotin carboxylase were always differentially expressed. The next enzymes in the cascade of the fatty acid synthesis were all up regulated, The complex of enzyme of the fatty acid synthase (*FAS*), the Beta-ketoacyl-ACP and Beta-ketoacyl-ACP synthase III and four other enzyme involved in the process of fatty acid elongation were all up-regulated. The *FatM* gene involved in elongation of fatty acid into palmitic acid was found in two copies in the cassava genome, both transcript being up regulated. Two other genes, *FatA* and *FatB*, common to all plants and involved in the same process of fatty acid elongation, were

transcribed but their transcription did not change (*FatB*) or poorly changed (*FatA*) when cassava plant was in symbiosis with *R. irregularis*. From this step of the fatty acid synthesis it was hypothesized (Bravo et *al.*, 2017) that during AMF symbiosis, palmitic acid is transferred to the endoplasmic reticulum with other fatty acids. It is first transformed in acyl-CoA with the help of an acyl-CoA synthetase. With the addition of glycerol-3-phosphate the previous product is transformed with the support of the mycorrhizal conserved enzyme RAM2 into β-MAG. β-MAG was hypothesized be then transferred fro the plant to the fungus via the STR/STR2 transporters. In the cassava, we did not detect any change in expression of the annotated acyl-CoA synthetase. However, this enzyme transformed any type of fatty acid into acyl-CoA, the palmitic acid produced by the *FatM* as well as the other form of fatty acid produced by *FatA* and *FatB*, such that the activity of this enzyme is probably not much influenced by an increase in the palmitic acid concentration. As suggested by Bravo et *al.*, 2017, the *RAM2* (GPAT) was up regulated as well as the *STR*, ABC transporters, mainly the *STR2*. β-MAG is, thus, possibility one form of fatty acid also transported from the cassava to the different *R. irregularis* isolates. However another cassava GPAT was also DE across all fungal treatments, this enzyme traditionally transforms the acyl-CoA into the first compound of the phospholipids, the lysophospholipidic acid. The detection of certain transporters and enzymes differentially expressed in the plant and in the fungi during AMF symbiosis might suggest that other ways could be involved in the transfer of fatty acids. Indeed, among the other transporters detected in the plant to be differentially express, we find seven ABC transporters, one being a conserved transporter across mycorrhizal plants (Delaux et *al.*, 2014) and two being hub genes in the blue plant module. More interestingly a plant phospholipid-transporting ATPase was among the hub gene of the blue plant module and was always DE with any of the *R. irregularis* groups. On the fungal side two phospholipid-transporting ATPase were differentially express and up regulated in GP1 compared to GP4. Both were high to highly correlated to the expression of the plant phospholipid-transporter. Phospholipid-transporters could transport phosphatidylcholine (PC), phosphatidyethanoamine (PE), phosphatidylglycerol (PG), phosphatidylinositol (PI), phosphatidic acid (PA) and phosphatidylserine (PS). PC and PE were found to be in lower concentration and were the only phospholipids changing their concentration in the mycorrhized mutant plant roots where the conserved gene module *FatM*, *RAM2* and *STR* were inactivated (Bravo et *al.*, 2017). Moreover, the three molecular species of PC (32:4, 36:4 and 36:5) with the highest concentration in mock roots, and that showed the strongest reduction in concentration when the roots were inoculated, are also the three highest PC molecular form found in ERM

(Wewer et *al.,* 2014). Lyso-phosphatidylcholine was reported to be a signal of the AMF fungus able to activate potato phosphate transporter genes (Drissner et *al.,* 2007). It is, thus, plausible that at least PC could be transferred from the host to the fungi and inversely. This assumption is supported on the fungus side by the fact that in the fungi, a phospholipase D nuclease found in the lightcyan module was differentially expressed between GP1 and GP4 and positively correlated with colonization. Phospholipase D hydrolysed the phosphatidylcholine into phosphatidic acid and choline. Moreover, a diacylglycerol cholinephosphotransferase was also found in the lightcyan module correlating to the colonizer type, this enzyme also transforms phoshpatidylcholine into choline and 1,2diacylglycerol (DAG). Both enzymes were correlated with each other. A low amount of DAG are found in the extra radical mycelium (ERM). However, their counterpart triacylglycerol (TAG) represented up to 90% of the total lipids. Thus we can suggest that DAG are rapidly transformed in the fungi into TAG non-polar storage lipids (Wewer et *al.,* 2014).

The phosphatidic acid obtained by phospholipase D hydrolyse of the phosphatidycholine, could be transformed into glycerol-3-phosphate by the action of a glycerol-3-phosphate O-acyltransferase. Glycerol-3-phosphate is the substrate for two other enzymes that we found differentially expressed between distant genetic groups. The first one, a glycerol kinase, uses the glycerol-3-phosphate to produce glycerol and inversely. The second one, a glycerol-3-phosphate deshydrogenase is a key enzyme at the crossroads between lipid metabolism and carbohydrate metabolism. This catalyses the reversible transformation of glycerol-3-phosphate into dihydroxacteone phosphate; a molecule that enters glycolysis. Wei et *al.,* (2004) have shown that disruption of this enzyme in the hemibiotrophic fungal plant pathogen *Colletotrichum gloeosporioides* resulted in severe defects in assimilation of glucose and amino acids as well as failure to form conidiat and resulted in arrhythmic growth. The only metabolite restoring the fungi was glycerol. Moreover, the defective fungal strain in glycerol-3-phosphate deshydrogenase performed normally and continued to grow. Finally, a reduction in glycerol was detected in the peripheral area around the fungi. Wei et *al.,* (2004) concluded that this hemibiotrophic fungus could survive only on glycerol transferred by the plant. It is also suggest that the genes involved in the biosynthesis of glycerol from glycolytic intermediates is one of the pathways that appears not to be conserved in obligate biotrophic pathogens (Tisserant et *al.,* 2012). If arbuscular fungi would rely on the glycerol as a carbon source we should find evidence that at low expression of glycerol-3-phosphate dehydrogenase there is an increase of glycerol in the fungi. The negative relationship between the expression of the glycerol kinase, transforming glycerol into glycerol-3-phosphate and the glycerol-3-

phosphate dehydrogenase, suggest that there is an increase transformation of the glycerol when the glycerol-3-phosphate dehydrogenase has low activity. Moreover, glycerol kinase is more active in low colonizers, in contrast to glycerol-3-phosphate dehydrogenase.

We therefore propose the hypothesis that arbuscular mycorrhizal fungi rely on the plant phosphatidylcholine as a carbon source to produce TAG, their major storage lipids.

We would also suggest that depending on the colonizer status of the isolate the phosphatidylcholine could be partially replace by glycerol as the source of carbon, such that low colonizer with low level of glycerol-3-phosphate dehydrogenase will have a high level of glycerol kinase, increasing the transformation of glycerol into glycerol-3-phosphate.

Ammonium transport

Finally, we found an intense ammonium transmembrane transport enriched, with five ammonium transporters differentially expressed in the plant. Such a finding is consistent with previous findings (Tisserant et *al.*, 2012), and with the common view of AMF transferring nitrogen to the plants. In parallel in the fungi, three glutamine synthetase were differentially expressed between genetic groups, suggesting a reduction or an increase in the synthesis of the glutamine depending on nitrogen transferred to the plant. Moreover, we found a nitrate transporter that varied with the different genetic groups as well as a nitrate reductase that was found in the lightcyan module, negatively related to the module and to colonization.

Conclusion

In conclusion, the diversity of isolates with their variability in gene expression was a good tool to unravel the important fungal and plant genes allowing symbiosis. Among them, we particularly underline the possible effect of a fungal chitin deacetylase able to mask the presence of the fungi to the chitin plant receptors, as well as the variability in expression of four central enzymes, a glycerol kinase, a glycerol-3-phosphatase, a phosphatase D and a diacylglycerol cholinephosphotransferase present in the fungal fatty acid pathway and related to the potential uptake of the phosphatidily choline and glycerol by the fungi.

Acknowledgements

We thank Nicolas Ruch for his precious help culturing plants. We thank also Aleš Látr at Symbiom, Amelia Campubri and Yolande Dalpé for allowing us to use their isolates. We would like also to thanks Consolée Alleti and Catherine Berney for their help and advice in the lab, Lucas Villard, Guillaume Cossard and Athimed El Taher for helpful comments and discussions on RNA-seq analysis and data interpretation. We thank Keith Harshman and all the team at the Lausanne University Genomic Technologies Facility. All bioinformatics computations were performed at the Vital-IT (www.vital-it.ch) Center for High Performance Computing of the Swiss Institute of Bioinformatics. This study was funded by the Swiss National Science Foundation (Grant number: 31003A_162549 to IRS).

References

Abouheif E. 1999. A method for testing the assumption of phylogenetic independence in comparative data. *Evolutionary Ecology Research* 1: 895-909.

Al-Karaki G, McMichael B, Zak J. 2004. Field response of wheat to arbuscular mycorrhizal fungi and drought stress. *Mycorrhiza* 14(4): 263-269.

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.

Angelard C, Sanders IR. 2011. Effect of segregation and genetic exchange on arbuscular mycorrhizal fungi in colonization of roots. *New Phytologist* 189(3): 652-657.

Antolín-Llovera M. Petutsching EK. Ried MK. Lipka V. Nürnberger T. Robatzek S. Parniske M. 2014. Knowing your friends and foes – plant receptor-like kinases as initiators of symbiosis or defence. *New Phytologist* 204: 791–802. doi:10.1111/nph.13117

Benjamini Y, Hochberg Y. 1995. CONTROLLING THE FALSE DISCOVERY RATE - A PRACTICAL AND POWERFUL APPROACH TO MULTIPLE TESTING. *Journal of the Royal Statistical Society Series B-Methodological* 57(1): 289-300.

Blomberg SP, Garland T, Ives AR. 2003. Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution* 57(4): 717-745.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15): 2114-2120.

Bravo, A. Brands M. Wewer V. Dörmann P. Harrison MJ. 2017. Arbuscular mycorrhiza-specific enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular mycorrhiza. *New Phytologist*. doi:10.1111/nph.14533

Bravo A, York T, Pumplin N, Mueller LA, Harrison MJ. 2016. Genes conserved for arbuscular mycorrhizal symbiosis identified through phylogenomics. *Nature Plants* 2(2).
Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* 34(5): 525-527.

Bredeson JV, Lyons JB, Prochnik SE, Wu GA, Ha CM, Edsinger-Gonzales E, Grimwood J, Schmutz J, Rabbi IY, Egesi C, et al. 2016. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nature Biotechnology* 34(5): 562-570.

Bruns TD, Taylor JW. 2016. FUNGAL SYMBIONTS Comment on "Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism". *Science* 351(6275).

Camilios-Neto D, Bonato P, Wassem R, Tadra-Sfeir MZ, Brusamarello-Santos LCC, Valdameri G, Donatti L, Faoro H, Weiss VA, Chubatsu LS, et al. 2014. Dual RNA-seq transcriptional analysis of wheat roots colonized by Azospirillum brasilense reveals up-regulation of nutrient acquisition and cell cycle genes. *Bmc Genomics* 15.

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

Ceballos I, Ruiz M, Fernandez C, Pena R, Rodriguez A, Sanders IR. 2013. The In Vitro Mass-Produced Model Mycorrhizal Fungus, Rhizophagus irregularis, Significantly Increases Yields of the Globally Important Food Security Crop Cassava. *Plos One* 8(8).

Cervantes-Gamez RG, Bueno-Ibarra MA, Cruz-Mendivil A, Calderon-Vazquez CL, Ramirez-Douriet CM, Maldonado-Mendoza IE, Villalobos-Lopez MA, Valdez-Ortiz A, Lopez-Meyer M. 2016. Arbuscular Mycorrhizal Symbiosis-Induced Expression Changes in Solanum lycopersicum Leaves Revealed by RNA-seq Analysis. *Plant Molecular Biology Reporter* 34(1): 89-102.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.

Bioinformatics 21(18): 3674-3676.
Cord-Landwehr S, Melcher RLJ, Kolkenbrock S, Moerschbacher BM. 2016. A chitin deacetylase from the endophytic fungus Pestalotiopsis sp efficiently inactivates the elicitor activity of chitin oligomers in rice cells. *Scientific Reports* 6.

Damron FH, Oglesby-Sherrouse AG, Wilks A, Barbier M. 2016. Dual-seq transcriptomics reveals the battle for iron during Pseudomonas aeruginosa acute murine pneumonia. *Scientific Reports* 6.

Das A, Saha D, Mondal TK. 2013. An optimized method for extraction of RNA from tea roots for functional genomics analysis. *Indian Journal of Biotechnology* 12(1): 129-132.

Davison J, Moora M, Ouml, pik M, Adholeya A, Ainsaar L, Ba A, Burla S, Diedhiou AG, Hiiesalu I, et al. 2015. Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science* 349(6251): 970-973.

Delaux PM, Varala K, Edger PP, Coruzzi GM, Pires JC, Ane JM. 2014. Comparative Phylogenomics Uncovers the Impact of Symbiotic Associations on Host Genome Evolution. *Plos Genetics* 10(7).

Dixon P. 2003. VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* 14(6): 927-930.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1): 15-21.
Evelin H, Kapoor R, Giri B. 2009. Arbuscular mycorrhizal fungi in alleviation of salt stress: a review. *Annals of Botany* 104(7): 1263-1280.

FAO. 2013. Howeler R, Lutaladio N, Thomas G. Cassava, a guide to sustainable production intensification. Rome.

Feng G, Zhang FS, Li XL, Tian CY, Tang C, Rengel Z. 2002. Improved tolerance of maize plants to salt stress by arbuscular mycorrhiza is related to higher accumulation of soluble sugars in roots. *Mycorrhiza* 12(4): 185-190.

Fiorilli V, Vallino M, Biselli C, Faccio A, Bagnaresi P, Bonfante P. 2015. Host and non-host roots in rice: cellular and molecular approaches reveal differential responses to arbuscular mycorrhizal fungi. *Frontiers in Plant Science* 6.

Frenzel A, Manthey K, Perlick AM, Meyer F, Puhler A, Kuster H, Krajinski F. 2005. Combined transcriptome profiling reveals a novel family of arbuscular mycorrhizal-specific Medicago truncatula lectin genes. *Molecular Plant-Microbe Interactions* 18(8): 771-782.

Gaiji N, Cardinale F, Prandi C, Bonfante P, Ranghino G. 2012. The computational-based structure of Dwarf14 provides evidence for its role as potential strigolactone receptor in plants. *BMC Research Notes* 5(1): 307

Galindo-Gonzalezi L, Deyholos MK. 2016. RNA-seq Transcriptome Response of Flax (Linum usitatissimum L.) to the Pathogenic Fungus Fusarium oxysporum f. sp lini. *Frontiers in Plant Science* 7.

Garcia K, Chasman D, Roy S, JM Ane. 20017. Physiological responses and gene co-expression network of mycorrhizal roots under K+ deprivation. *Plant Physiology* 173(4):

Govindarajulu M, Pfeffer PE, Jin HR, Abubaker J, Douds DD, Allen JW, Bucking H, Lammers PJ, Shachar-Hill Y. 2005. Nitrogen transfer in the arbuscular mycorrhizal symbiosis. *Nature* 435(7043): 819-823.

Gutjahr C, Radovanovic D, Geoffroy J, Zhang Q, Siegler H, Chiapello M, Casieri L, An K, An G, Guiderdoni E, et al. 2012. The half-size ABC transporters STR1 and STR2 are indispensable for mycorrhizal arbuscule formation in rice. *Plant Journal* 69(5): 906-920.

Gutjahr C, Sawers RJH, Marti G, Andres-Hernandez L, Yang SY, Casieri L, Angliker H, Oakeley EJ, Wolfender JL, Abreu-Goodger C, et al. 2015. Transcriptome diversity among rice root types during asymbiosis and interaction with arbuscular mycorrhizal fungi. *Proceedings of the National Academy of Sciences of the United States of America* 112(21): 6754-6759.

Hajiboland R, Aliasgharzadeh N, Laiegh SF, Poschenrieder C. 2010. Colonization with arbuscular mycorrhizal fungi improves salinity tolerance of tomato (Solanum lycopersicum L.) plants. *Plant and Soil* 331(1-2): 313-327.

Handa Y, Nishide H, Takeda N, Suzuki Y, Kawaguchi M, Saito K. 2015. RNA-seq Transcriptional Profiling of an Arbuscular Mycorrhiza Provides Insights into Regulated and Coordinated Gene Expression in Lotus japonicus and Rhizophagus irregularis. *Plant and Cell Physiology* 56(8): 1490-1511.

Kamel L, Keller-Pearson M, Roux C, Ane JM. 2017. Biology and evolution of arbuscular mycorrhizal symbiosis in the light of genomics. *New Phytologist* 213(2): 531-536.

Karandashov V, Bucher M. 2005. Symbiotic phosphate transport in arbuscular mycorrhizas. *Trends in Plant Science* 10(1): 22-29.

Keck F, Rimet F, Bouchez A, Franc A. 2016. phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecology and Evolution* 6(9): 2774-2780.

Kloppholz S, Kuhn H, Requena N. 2011. A Secreted Fungal Effector of Glomus intraradices Promotes Symbiotic Biotrophy. *Current Biology* 21(14): 1204-1209.

Koch AM, Croll D, Sanders IR. 2006. Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. *Ecology Letters* 9(2): 103-110.

Koch AM, Kuhn G, Fontanillas P, Fumagalli L, Goudet J, Sanders IR. 2004. High genetic variability and low local diversity in a population of arbuscular mycorrhizal fungi. *Proceedings of the National Academy of Sciences of the United States of America* 101(8): 2369-2374.

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *Bmc Bioinformatics* 9.

Li NN, Xu CC, Li-Beisson YH, Philippar K. 2016. Fatty Acid and Lipid Transport in Plant Cells. *Trends in Plant Science* 21(2): 145-158.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30(7): 923-930.

Lin K, Limpens E, Zhang ZH, Ivanov S, Saunders DGO, Mu DS, Pang EL, Cao HF, Cha HH, Lin T, et al. 2014. Single Nucleus Genome Sequencing Reveals High Similarity among Nuclei of an Endomycorrhizal Fungus. *Plos Genetics* 10(1).

Liu HZ, Wang XE, Zhang HJ, Yang YY, Ge XC, Song FM. 2008. A rice serine carboxypeptidase-like gene OsBISCPL1 is involved in regulation of defense responses against biotic and oxidative stress. *Gene* 420(1): 57-65.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15(12).

Mensah JA, Koch AM, Antunes PM, Kiers ET, Hart M, Bucking H. 2015. High functional diversity within species of arbuscular mycorrhizal fungi is associated with differences in phosphate and nitrogen uptake and fungal phosphate metabolism. *Mycorrhiza* 25(7): 533-546.

Moran PAP. 1948. Rank correlation and permutation distributions. *Proceedings of the Cambridge Philosophical Society* 44: 142-144.

Munkvold L, Kjoller R, Vestberg M, Rosendahl S, Jakobsen I. 2004. High functional diversity within species of arbuscular mycorrhizal fungi. *New Phytologist* 164(2): 357-364.

Niu XW, Zhao XQ, Ling KS, Levi A, Sun YY, Fan M. 2016. The FonSIX6 gene acts as an avirulence effector in the Fusarium oxysporum f. sp niveum - watermelon pathosystem. *Scientific Reports* 6.

Nivedita, Verma PK, Upadhyaya KC. 2017. Lectin Protein Kinase Is Induced in Plant Roots in Responseto the Endophytic Fungus, Piriformospora indica. *Plant Molecular Biology Reporter* doi:10.1007/s11105-017-1024-4

Nuss AM, Beckstette M, Pimenova M, Schmuhl C, Opitz W, Pisano F, Heroven AK, Dersch

P. 2017. Tissue dual RNA-seq allows fast discovery of infection-specific functions and riboregulators shaping host-pathogen transcriptomes. *Proceedings of the National Academy of Sciences of the United States of America* 114(5): E791-E800.

Opik M, Davison J, Moora M, Partel M, Zobel M. 2016. FUNGAL SYMBIONTS Response to Comment on "Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism". *Science* 351(6275).

Pagel M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology* 48: 612-622.

Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20(2): 289-290.

Parniske M. 2008. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nature Reviews Microbiology* 6(10): 763-775.

Pozo MJ, Azcon-Aguilar C. 2007. Unraveling mycorrhiza-induced resistance. *Current Opinion in Plant Biology* 10(4): 393-398.

Rodriguez A, Sanders IR. 2015. The role of community and population ecology in applying mycorrhizal fungi for improved food security. *Isme Journal* 9(5): 1053-1061.

Rooney HCE, van 't Klooster JW, van der Hoorn RAL, Joosten M, Jones JDG, de Wit P. 2005. Cladosporium Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science* 308(5729): 1783-1786.

Ropars J, Toro KS, Noel J, Pelin A, Charron P, Farinelli L, Marton T, Kruger M, Fuchs J, Brachmann A, et al. 2016. Evidence for the sexual origin of heterokaryosis in arbuscular mycorrhizal fungi. *Nature Microbiology* 1(6).

Sanchez-Garcia A, Moreno-Perez AJ, Muro-Pastor AM, Salas JJ, Garces R, Martinez-Force E. 2010. Acyl-ACP thioesterases from castor (Ricinus communis L.): An enzymatic system appropriate for high rates of oil synthesis and accumulation. *Phytochemistry* 71(8-9): 860-

869.
Sanders IR. 2010. 'Designer' mycorrhizas?: Using natural genetic variation in AM fungi to increase plant growth. *Isme Journal* 4(9): 1081-1083.

Santana MA, Romay G, Matehus J, Vicente-Villardon JL, Demey JR. 2009. A simple and low-cost strategy for micropropagation of cassava (Manihot esculenta Crantz). *African Journal of Biotechnology* 8(16): 3789-3797.

Savary R, Masclaux FG, Wyss T, Droh G, Cruz Corella J, Machado AP, Morton JB Sanders IR. Submitted. A population genomics approach shows widespread geographical distribution of cryptic genomic forms of the symbiotic fungus *Rhizophagus irregularis*. submitted to ISME journal 2017.

Shu B, Li WC, Liu LQ, Wei YZ, Shi SY. 2016. Transcriptomes of Arbuscular Mycorrhizal Fungi and Litchi Host Interaction after Tree Girdling. *Frontiers in Microbiology* 7.

Soneson C, Love MI, Robinson MD. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences [version 1; referees: 2 approved] *F1000Research* 4:1521
Song FQ, Li JZ, Fan XX, Zhang Q, Chang W, Yang FS, Geng G. 2016. Transcriptome analysis of Glomus mosseae/Medicago sativa mycorrhiza on atrazine stress. *Scientific Reports* 6.

Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J, Szczyglowski K, et al. 2002. A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* 417(6892): 959-962.

Sugimura Y, Saito K. 2017. Transcriptional profiling of arbuscular mycorrhizal roots exposed to high levels of phosphate reveals the repression of cell cycle-related genes and secreted protein genes in Rhizophagus irregularis. *Mycorrhiza* 27(2): 139-146.

Tisserant E, Kohler A, Dozolme-Seddas P, Balestrini R, Benabdellah K, Colard A, Croll D, Da Silva C, Gomez SK, Koul R, et al. 2012. The transcriptome of the arbuscular mycorrhizal fungus Glomus intraradices (DAOM 197198) reveals functional tradeoffs in an obligate

symbiont. *New Phytologist* 193(3): 755-769.
Tisserant E, Malbreil M, Kuo A, Kohler A, Symeonidi A, Balestrini R, Charron P, Duensing N, Frey NFD, Gianinazzi-Pearson V, et al. 2013. Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* 110(50): 20117-20122.

Toro KS, Brachmann A. 2016. The effector candidate repertoire of the arbuscular mycorrhizal fungus Rhizophagus clarus. *Bmc Genomics* 17.

Tsuzuki S, Handa Y, Takeda N, Kawaguchi M. 2016. Strigolactone-Induced Putative Secreted Protein 1 Is Required for the Establishment of Symbiosis by the Arbuscular Mycorrhizal Fungus Rhizophagus irregularis. *Molecular Plant-Microbe Interactions* 29(4): 277-286.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T, Wiemken A, Sanders IR. 1998. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* 396(6706): 69-72.

van der Heijden MGA, Martin FM, Selosse MA, Sanders IR. 2015. Mycorrhizal ecology and evolution: the past, the present, and the future. *New Phytologist* 205(4): 1406-1423.

Van der Hoorn RAL, Roth R, De Wit PJG. 2001. Identification of distinct specificity determinants in resistance protein Cf-4 allows construction of a Cf-9 mutant that confers recognition of avirulence protein AVR4. *Plant Cell* 13(2): 273-285.

Wan JR, Zhang XC, Neece D, Ramonell KM, Clough S, Kim SY, Stacey MG, Stacey G. 2008. A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. *Plant Cell* 20(2): 471-481.

Wang ET, Schornack S, Marsh JF, Gobbato E, Schwessinger B, Eastmond P, Schultze M, Kamoun S, Oldroyd GED. 2012. A Common Signaling Process that Promotes Mycorrhizal and Oomycete Colonization of Plants. *Current Biology* 22(23): 2242-2246.

Wei YD, Shen WY, Dauk M, Wang F, Selvaraj G, Zou JT. 2004. Targeted gene disruption of

glycerol-3-phosphate dehydrogenase in Colletotrichum gloeosporioides reveals evidence that glycerol is a significant transferred nutrient from host plant to fungal pathogen. *Journal of Biological Chemistry* 279(1): 429-435.

Westermann AJ, Forstner KU, Amman F, Barquist L, Chao YJ, Schulte LN, Muller L, Reinhardt R, Stadler PF, Vogel J. 2016. Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. *Nature* 529(7587): 496-+.

Westermann AJ, Gorski SA, Vogel J. 2012. Dual RNA-seq of pathogen and host. *Nature Reviews Microbiology* 10(9): 618-630.

Wewer V, Brands M, Dormann P. 2014. Fatty acid synthesis and lipid metabolism in the obligate biotrophic fungus Rhizophagus irregularis during mycorrhization of Lotus japonicus. *Plant Journal* 79(3): 398-412.

Wilson MC, Mutka AM, Hummel AW, Berry J, Chauhan RD, Vijayaraghavan A, Taylor NJ, Voytas DF, Chitwood DH, Bart RS. 2017. Gene expression atlas for the food security crop cassava. *New Phytologist* 213(4): 1632-1641.

Wyss T, Masclaux FG, Rosikiewicz P, Pagni M, Sanders IR. 2016. Population genomics reveals that within-fungus polymorphism is common and maintained in populations of the mycorrhizal fungus Rhizophagus irregularis. *Isme Journal* 10(10): 2514-2526.

Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* 11(2).

Zhang Q, Blaylock LA, Harrison MJ. 2010. Two Medicago truncatula Half-ABC Transporters Are Essential for Arbuscule Development in Arbuscular Mycorrhizal Symbiosis. *Plant Cell* 22(5): 1483-1497.

Zouari I, Salvioli A, Chialva M, Novero M, Miozzi L, Tenore GC, Bagnaresi P, Bonfante P. 2014. From root to fruit: RNA-Seq analysis shows that arbuscular mycorrhizal symbiosis may affect tomato fruit metabolism. *Bmc Genomics* 15.

Supplementary information

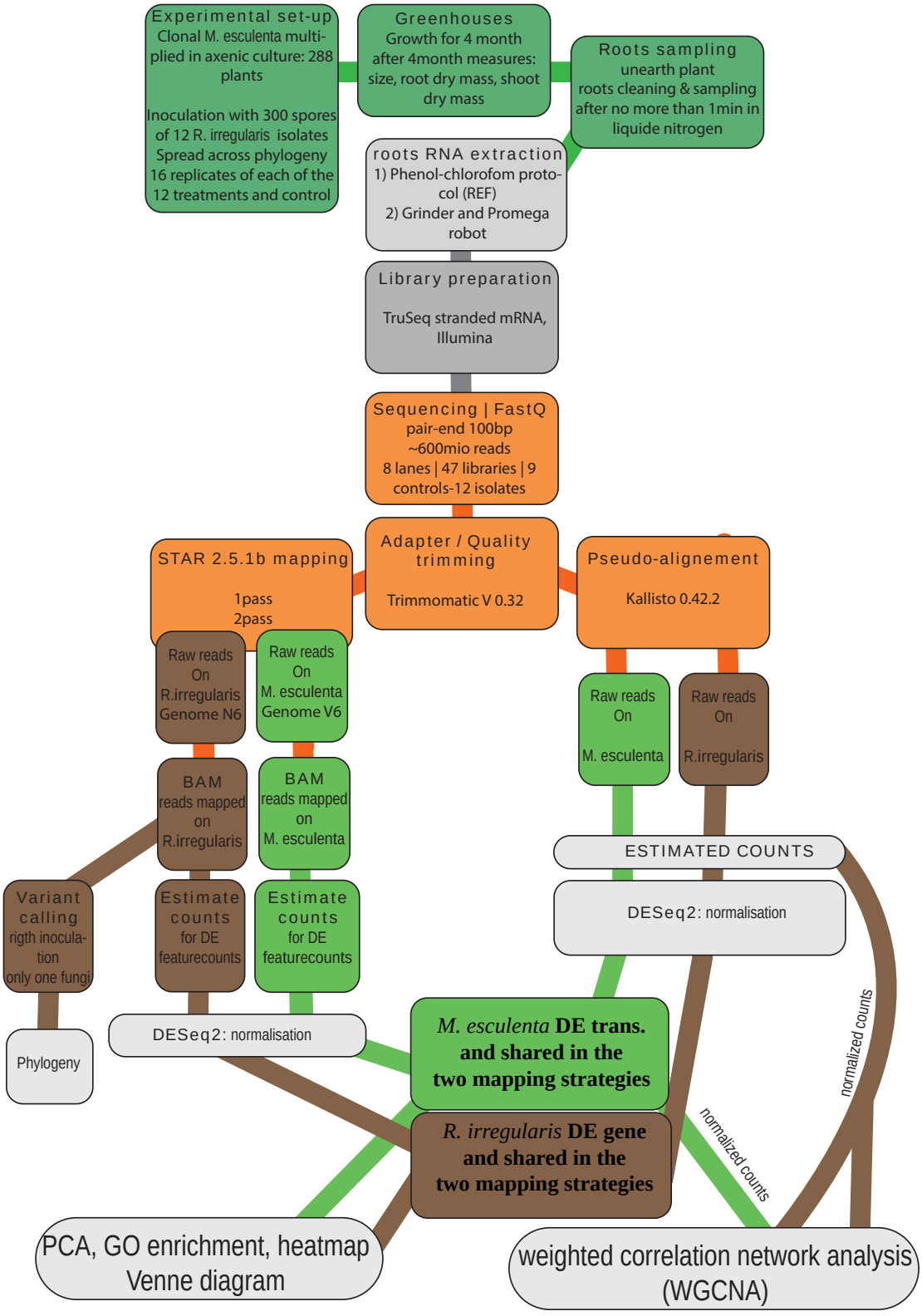


Fig. S1 The pipeline, from the experimental setup to the bioinformatic analysis of the expression

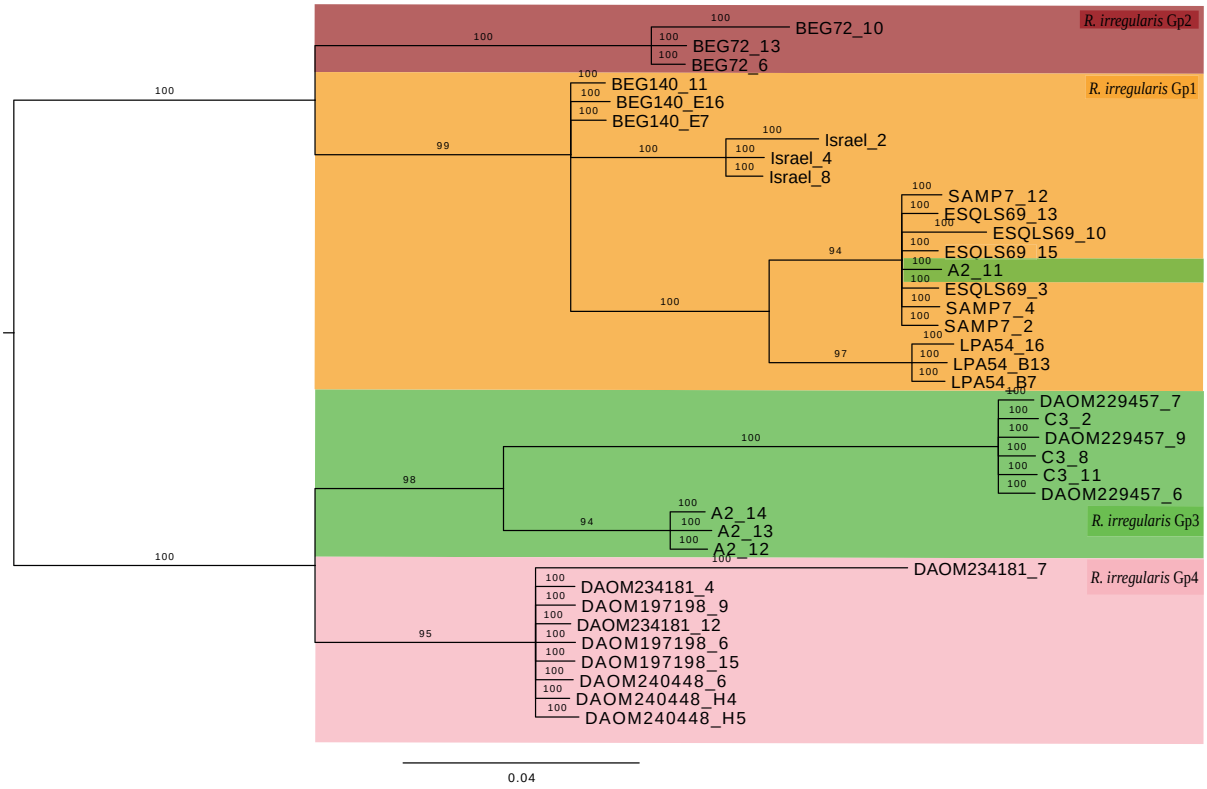


Fig. S2 Tree base on the concatenation of the *R. irregularis* called SNPs for each isolate and each replicates (n=38), in order to confirm the presence of the right fungal treatment in each pot. The treatment A2_11 represented a treatment supposed to have the isolate A2, which should have been clustering with other A2 treatments. A2_11 was then discared from the analysis.

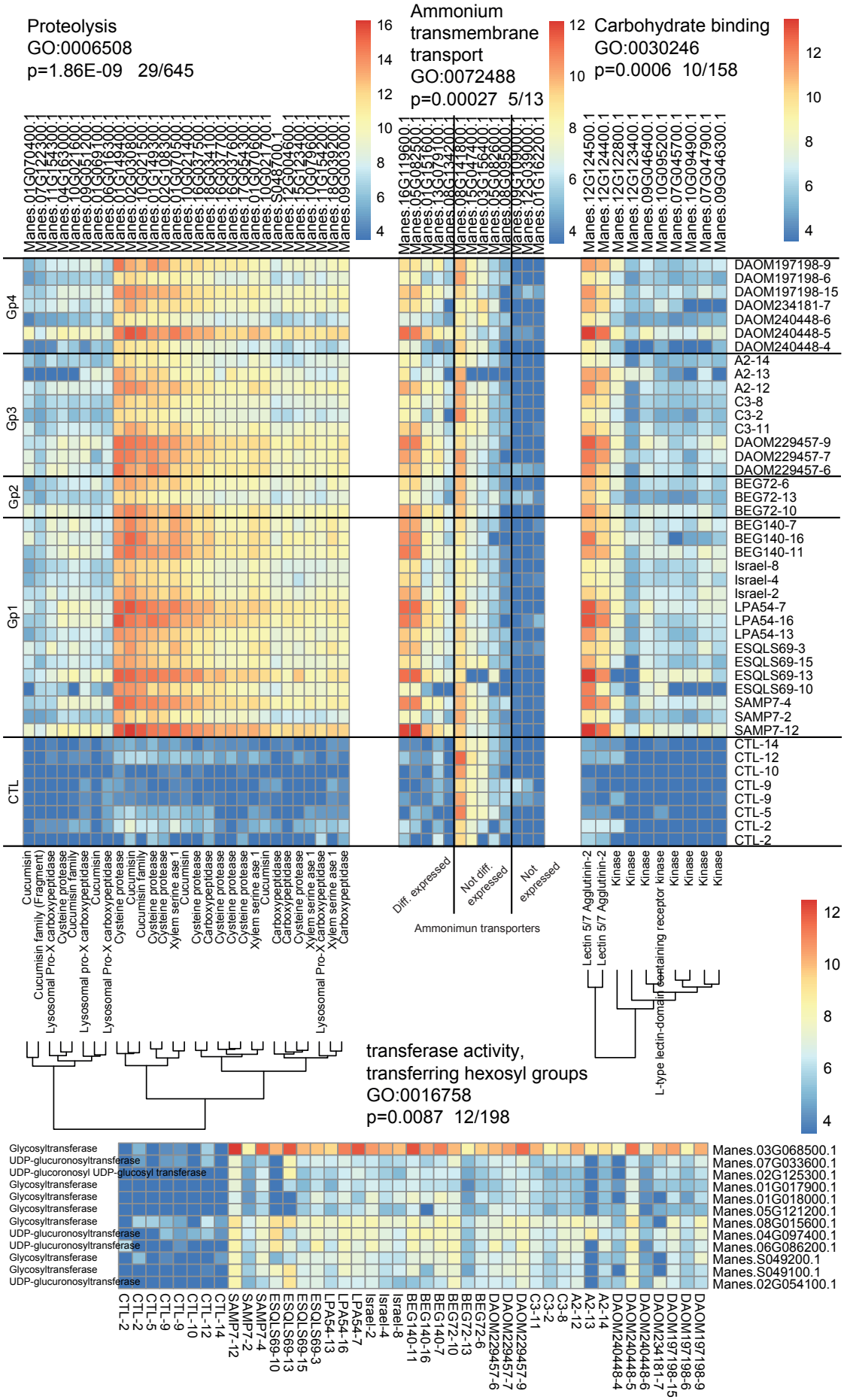

Proteolysis
GO:0006508
p=1.86E-09 29/645
Ammonium transmembrane transport
GO:0072488
p=0.00027 5/13
Carbohydrate binding
GO:0030246
p=0.0006 10/158
Gp4
Gp3
Gp2
Gp1
CTL
Diff. expressed
Not diff. expressed
Not expressed
Ammonimun transporters
transferase activity, transferring hexosyl groups
GO:0016758
p=0.0087 12/198

Fig. S3 Heatmaps of the normalized expression (Kallisto) of the cassava transcripts for all treatments, included in GO terms significantly enriched after p-value adjustement. The heatmap for the GO term proteolysis includes all the genes present in the GO enriched terms serine-type endopeptidase activity and cysteine-type peptidase activity. The genes for the GO enrich terms ammonium transmembrane transport, ammonium transmembrane transporter activity, carbohydrate binding, transferase activity, transferring hexosyl groups are also presented.

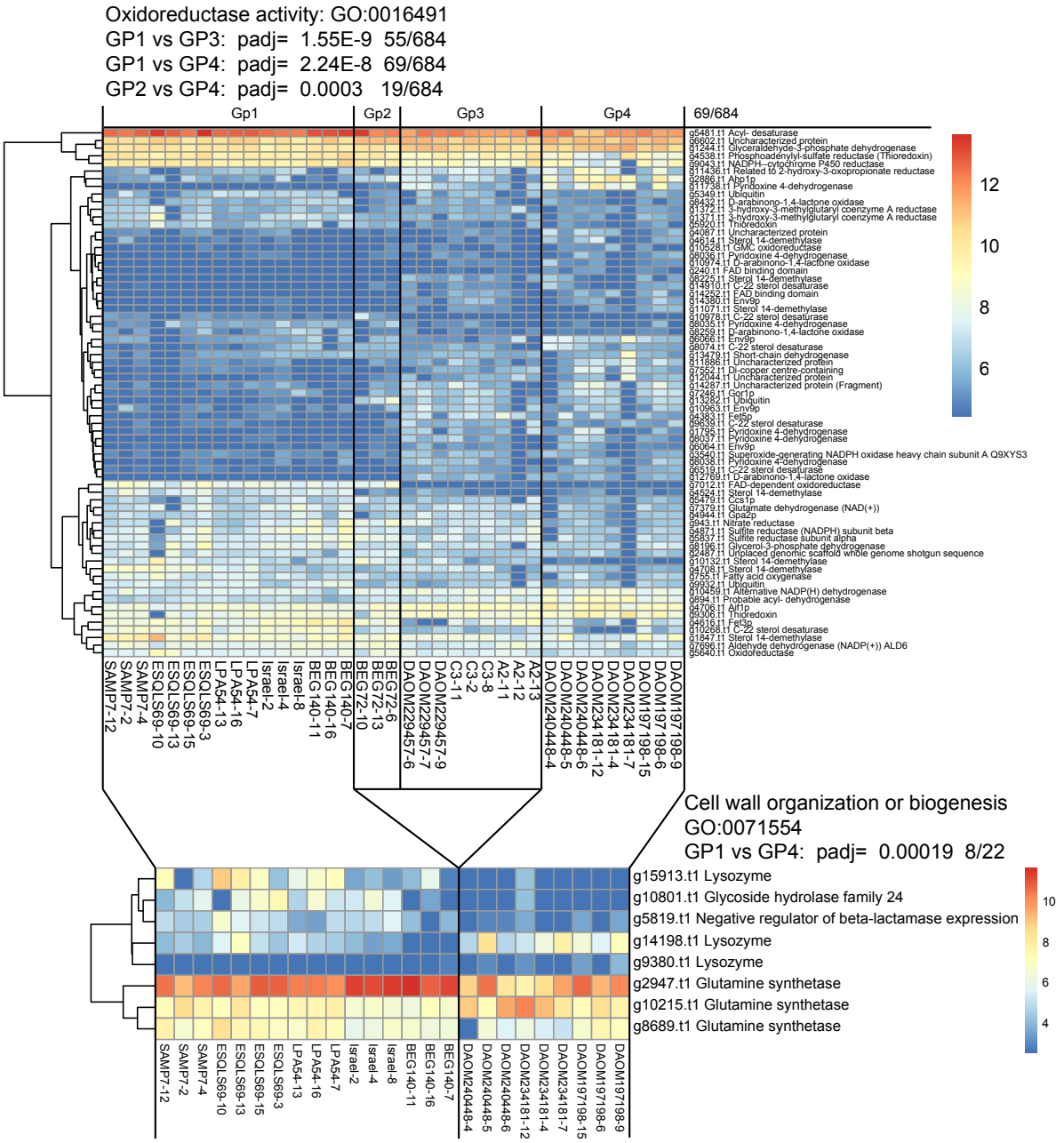


Fig. S4 Heatmaps of the normalized expression (Kallisto) of *R. irregularis* genes included in GO terms significantly enriched after p-value adjustement. The first heatmap, represent the genes differentially express for the oxidoreductase activity GO term significant between GP1 and GP3 and 4 and between GP2 and GP4. The second heatmap represent the cell wall organization or biogenesis GO term enrich in the DE genes between GP1 and Gp4

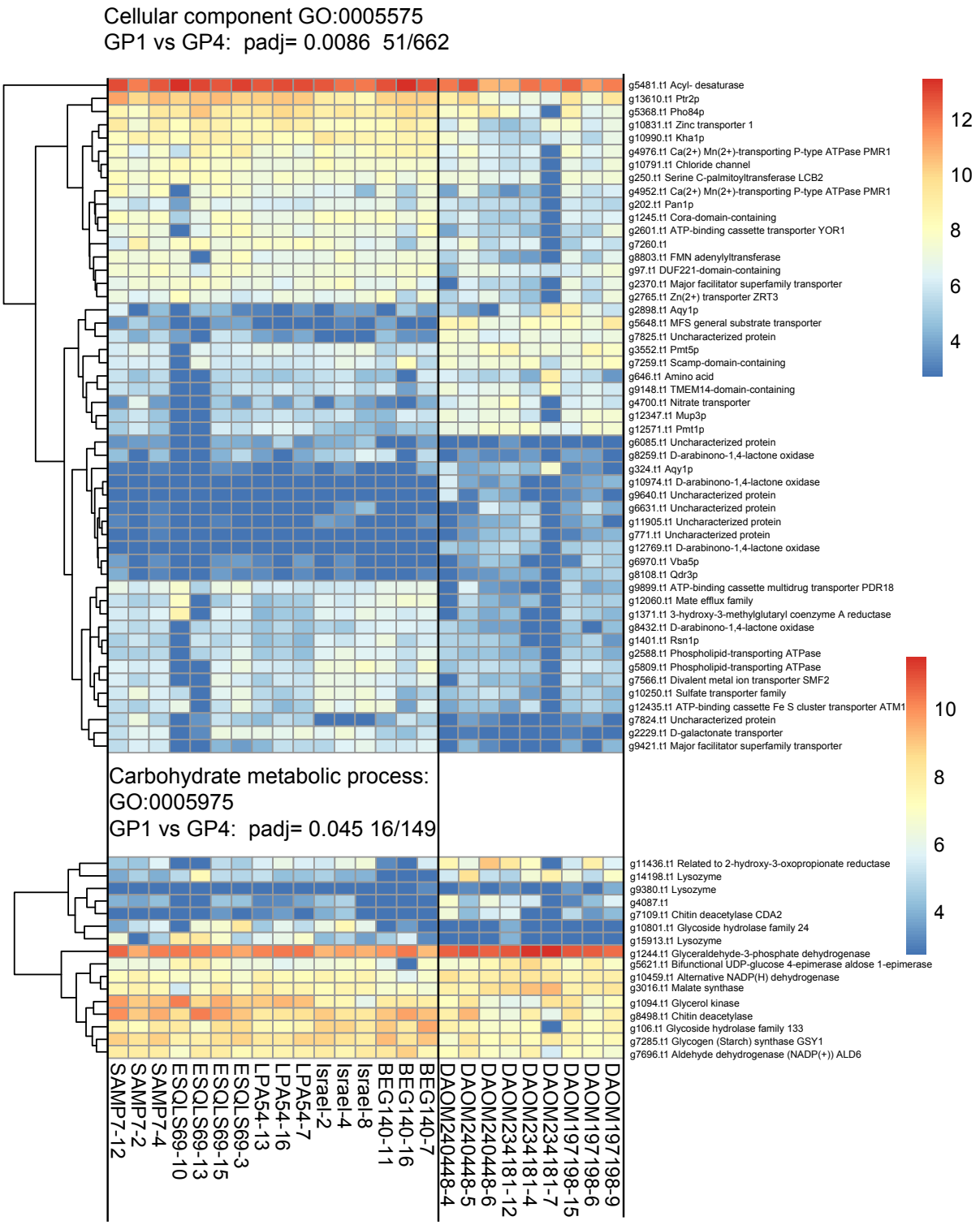


Fig. S5 Heatmaps of the normalized expression (Kallisto) of the *R. irregularis* genes included in GO terms significantly enriched after p-value adjustement for gene DE between GP1 and GP4. The first heatmap indicates the genes of the cellular component GO term. The second one, represent the gene of the Carbohydrate metabolic process GO term.

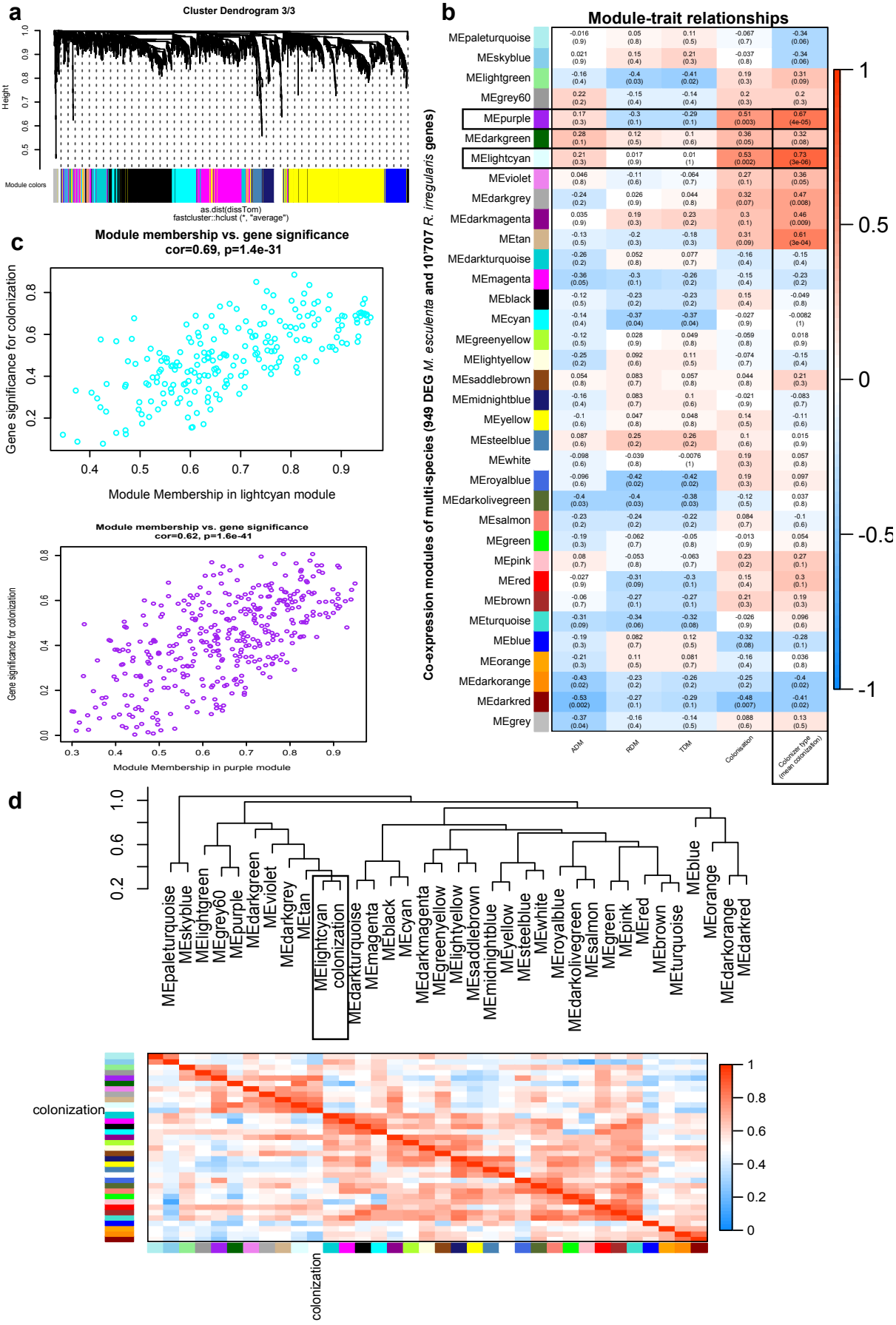

a
Cluster Dendrogram 3/3
Height
Module colors
as.dist(dissTom)
fastcluster::hclust (*, "average")
b
Module-trait relationships
Co-expression modules of multi-species (949 DEG M. esculenta and 10'707 R. irregularis genes)
MEpaleturquoise
MEskyblue
MElightgreen
MEgrey60
MEpurple
MEdarkgreen
MElightcyan
MEviolet
MEdarkgrey
MEdarkmagenta
MEtan
MEdarkturquoise
MEmagenta
MEblack
MEcyan
MEgreenyellow
MElightyellow
MEsaddlebrown
MEwhite
MEroyalblue
MEdarkolivegreen
MEsalmon
MEgreen
MEpink
MEred
MEbrown
MEturquoise
MEblue
MEorange
MEdarkorange
MEdarkred
MEgrey
ADM
RDM
TDM
Colonisation
Colonizer type (mean colonization)
c
Module membership vs. gene significance
cor=0.69, p=1.4e-31
Gene significance for colonization
Module Membership in lightcyan module
Module membership vs. gene significance
cor=0.62, p=1.6e-41
Gene significance for colonization
Module Membership in purple module
d
colonization

Fig. S6 The WGCNA multiple-species analysis, including the normalized expression (Kallisto) of the 949 DE *M. esculenta* transcript in the inoculated plants and the normalized expression of the 10’727 express fungal genes. (a) Part 3 of 3 of the cluster dendrogram of the gene in different modules. (b) Pearson correlations and p-value between module eigengene values and phenotypic values (ADM, RDM, TDM, Colonization, Colonizer type). The two most important modules are the lightcyan and the purple. (c) Relationship between module membership of a gene and gene significance for colonizer type. (d) Hierarchical clustering of eigengenes values of each modules and colonizer type.

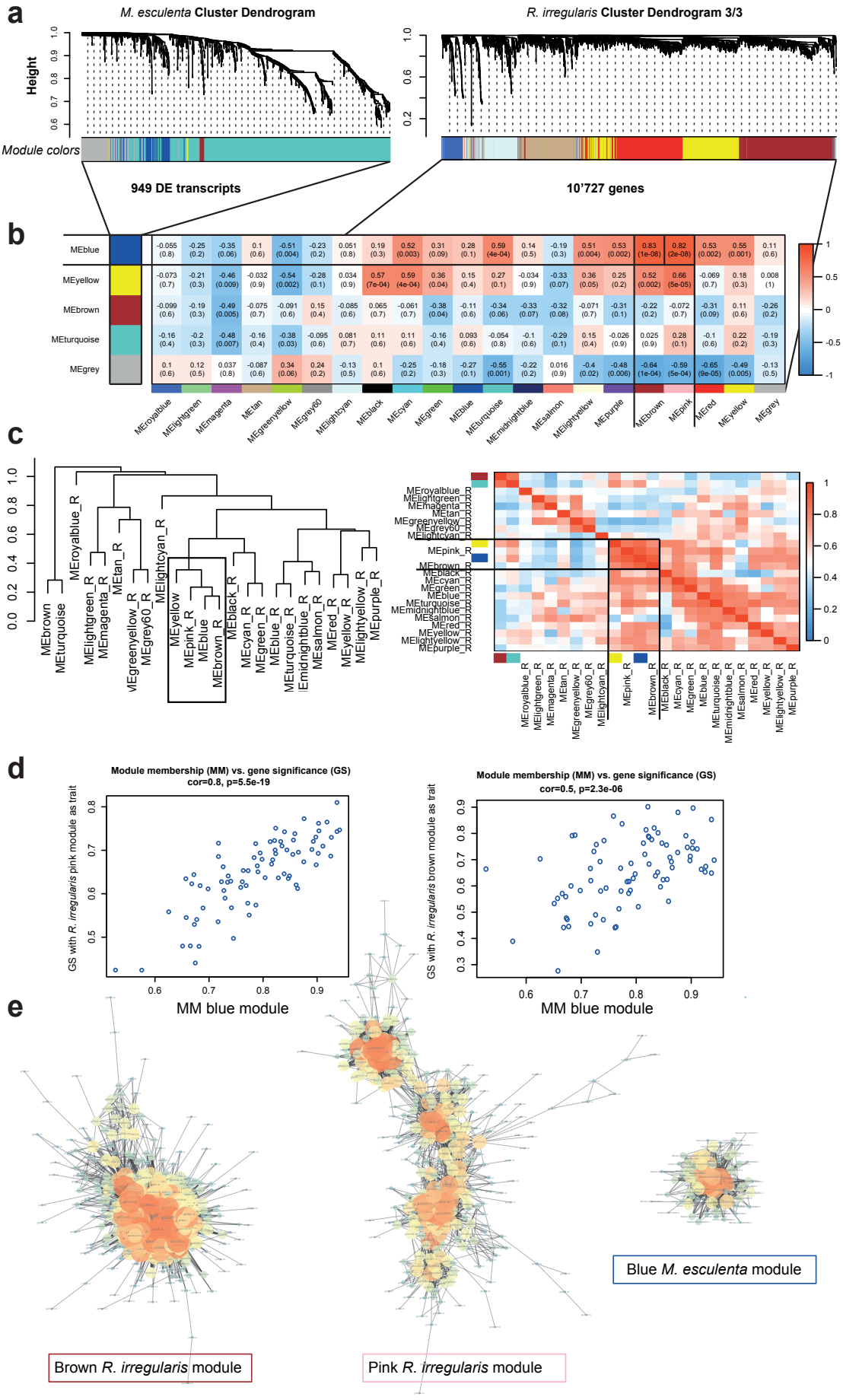

a
M. esculenta Cluster Dendrogram
R. irregularis Cluster Dendrogram 3/3
Height
Module colors
949 DE transcripts
10'727 genes
b
MEblue
MEyellow
MEbrown
MEturquoise
MEgrey
c
d
Module membership (MM) vs. gene significance (GS)
cor=0.8, p=5.5e-19
GS with R. irregularis pink module as trait
MM blue module
Module membership (MM) vs. gene significance (GS)
cor=0.5, p=2.3e-06
GS with R. irregularis brown module as trait
MM blue module
e
Brown R. irregularis module
Pink R. irregularis module
Blue M. esculenta module

Fig. S7 The second WGCNA analysis, (a) in parallel we performed the clustering of transcript/genes in modules of the normalized expression (Kallisto) of the 949 DE *M. esculenta* transcripts and of the normalized expression of the 10’727 express fungal genes. (b) Pearson correlations and p-value between plants modules eigengenes values and fungal modules eigengenes values. (c) Hierarchical clustering of eigengenes values of each plant and fungal modules. (d) Relationship between plant module membership of a gene and gene significance for both fungal gene modules. (e) Network of the blue plant gene module and the pink and brown fungal gene modules. Size and colors of the rounds surrounding each gene depict the centrality (degree) of the gene in the network.

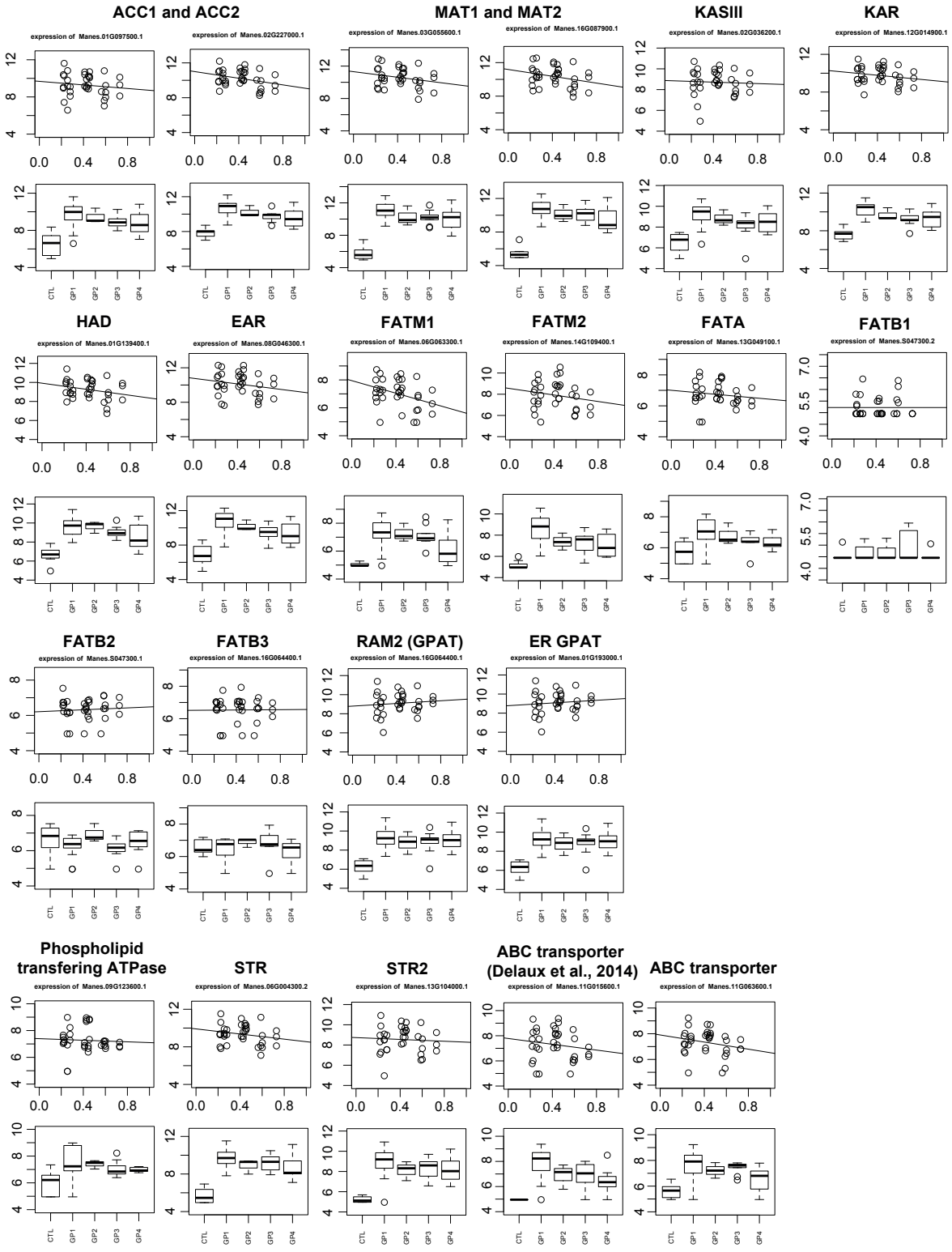


Fig. S8 Plot of the *M. esculenta* normalized expression (Kallisto) versus colonizer type and boxplot of normalized expression (Kallisto) for the CTL and each genetic group for each cassava transcript involved in the fatty acid biosynthesis pathway represented as a graphical schema in Fig 5.

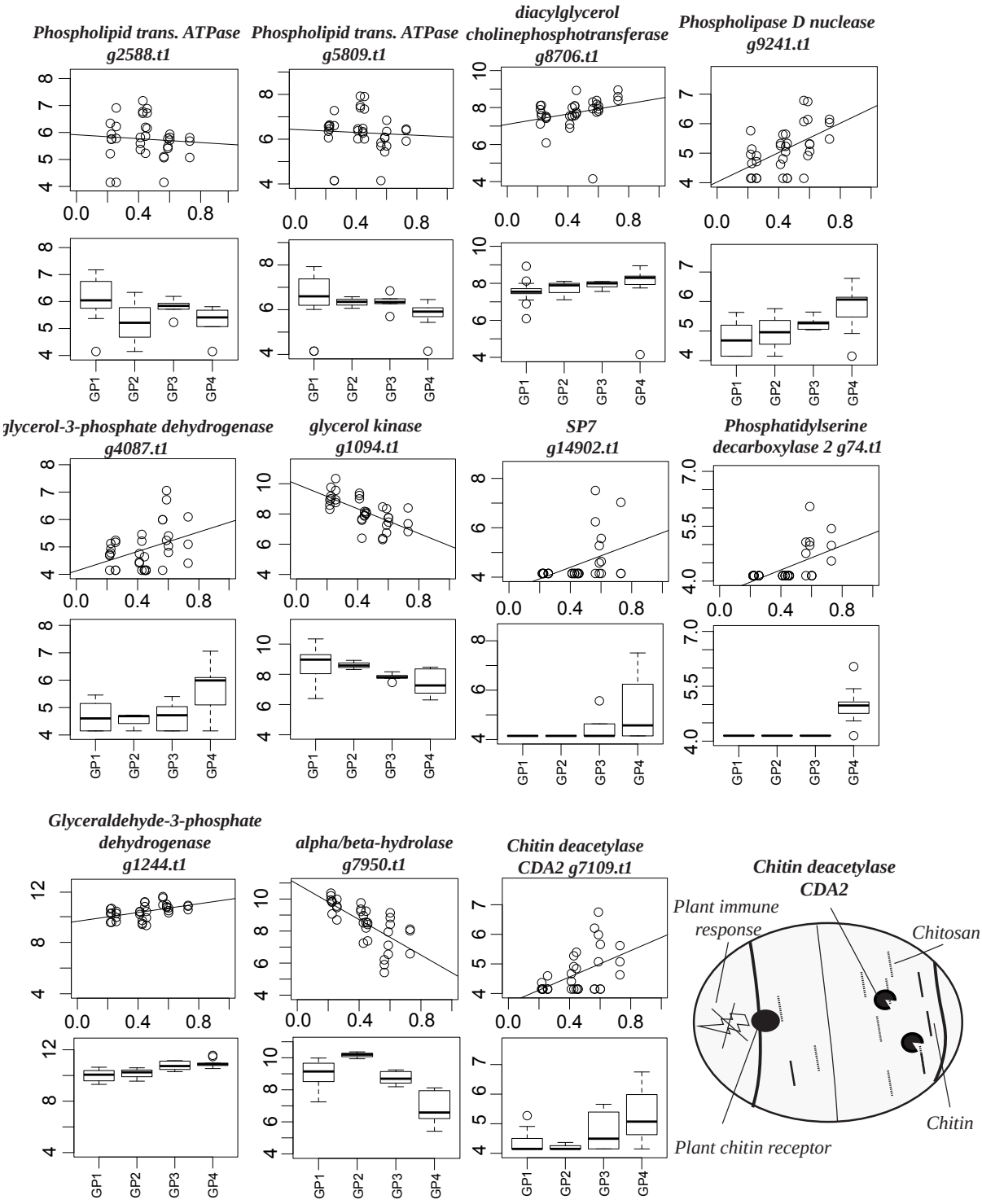


Fig. S9 Plot of the *R. irregularis* normalized expression (Kallisto) versus colonizer type and boxplot of normalized expression (Kallisto) for each genetic group for each fungal gene involved in the fatty acid biosynthesis pathway represented as a graphical schema in Fig 5.

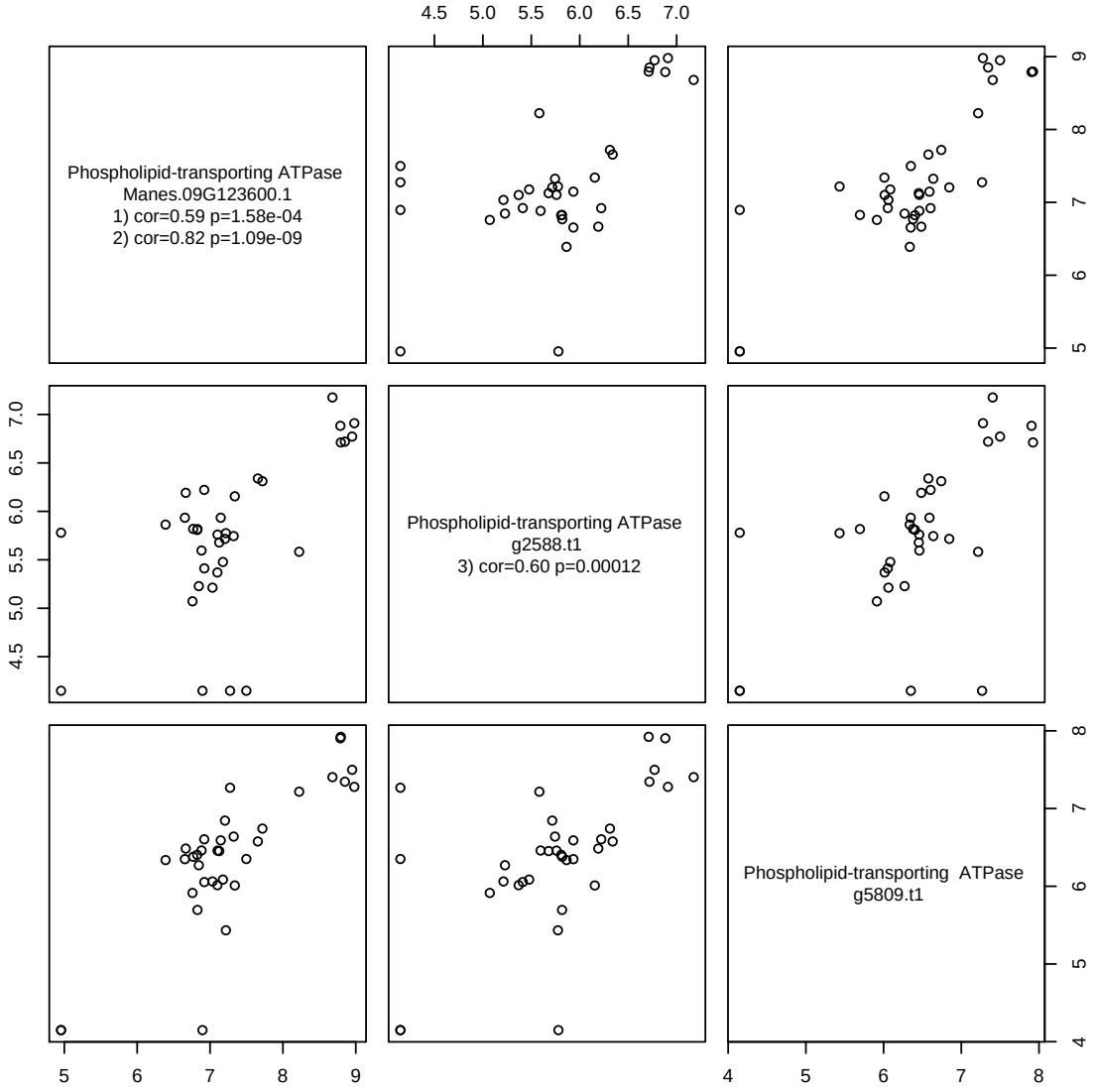


Fig. S10 Pearson between the plant phospholipid transporting ATPase and the two fungal phospholipid-transporting ATPase.

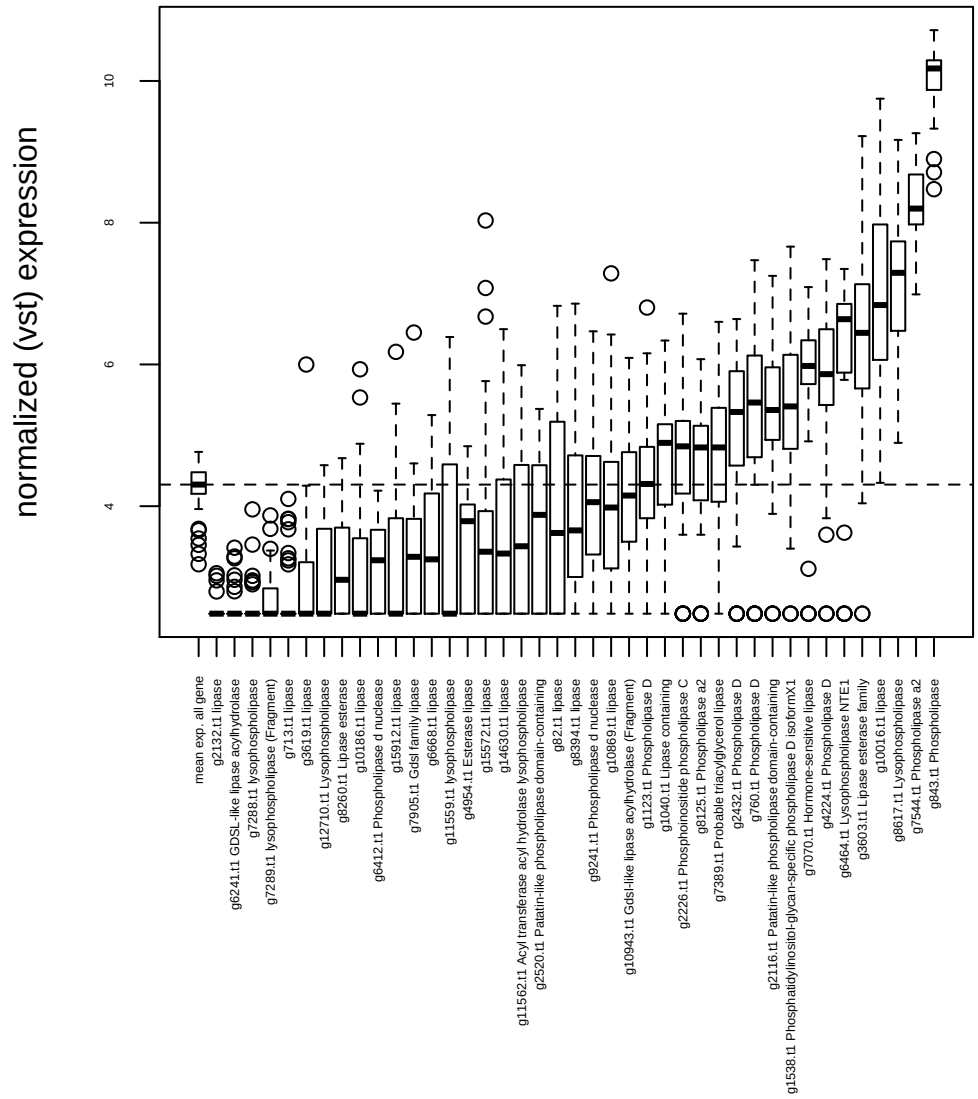


Fig. S11 Normalized expression of the annotated *R. irregularis* lipases.

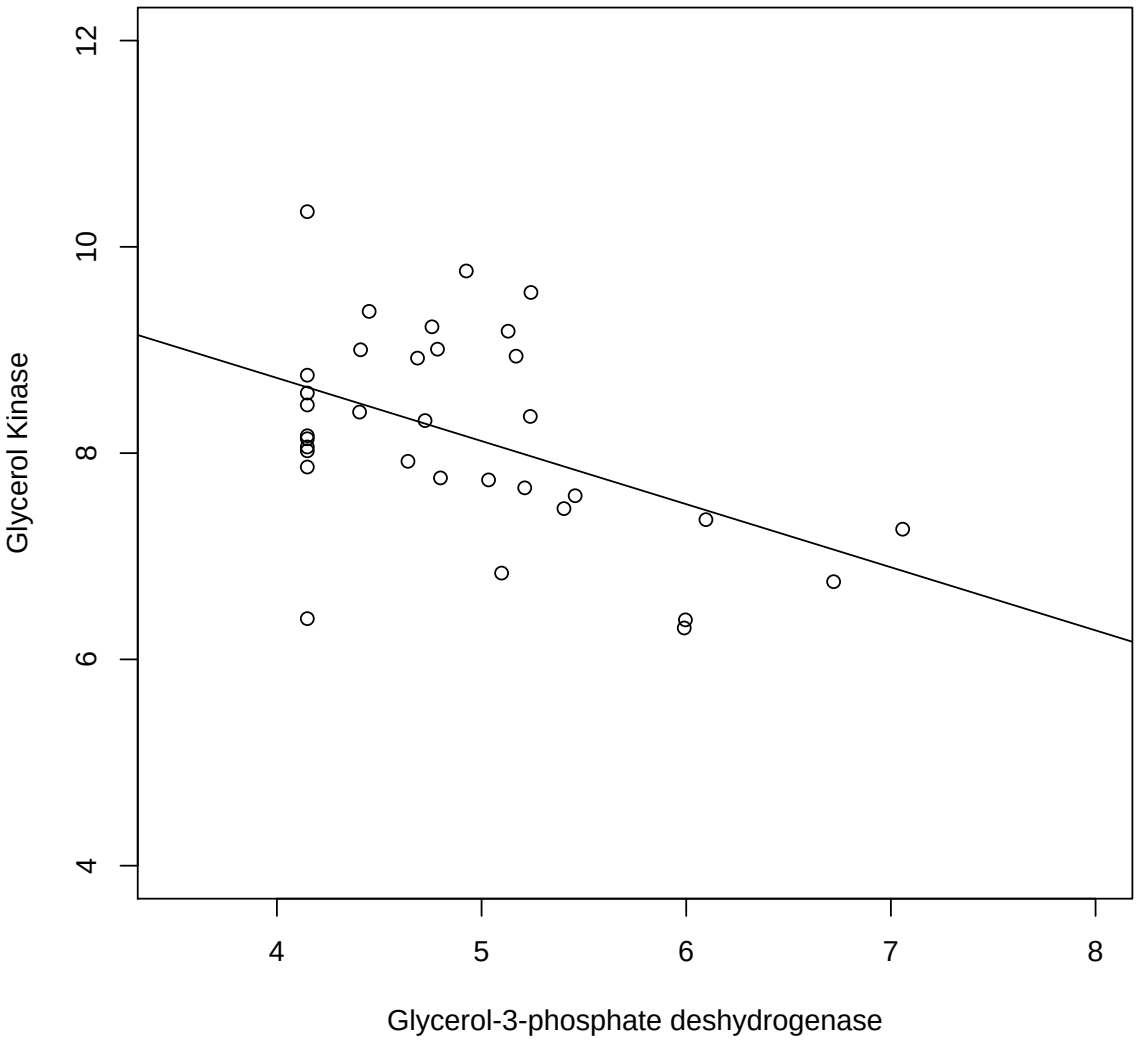


Fig. S12 Person correlation between the normalized expressions of the two fungal enzymes, the glycerol kinase and the glycerol-3-phosphate dehydrogenase using the glycerol-3-phosphate as substrate.

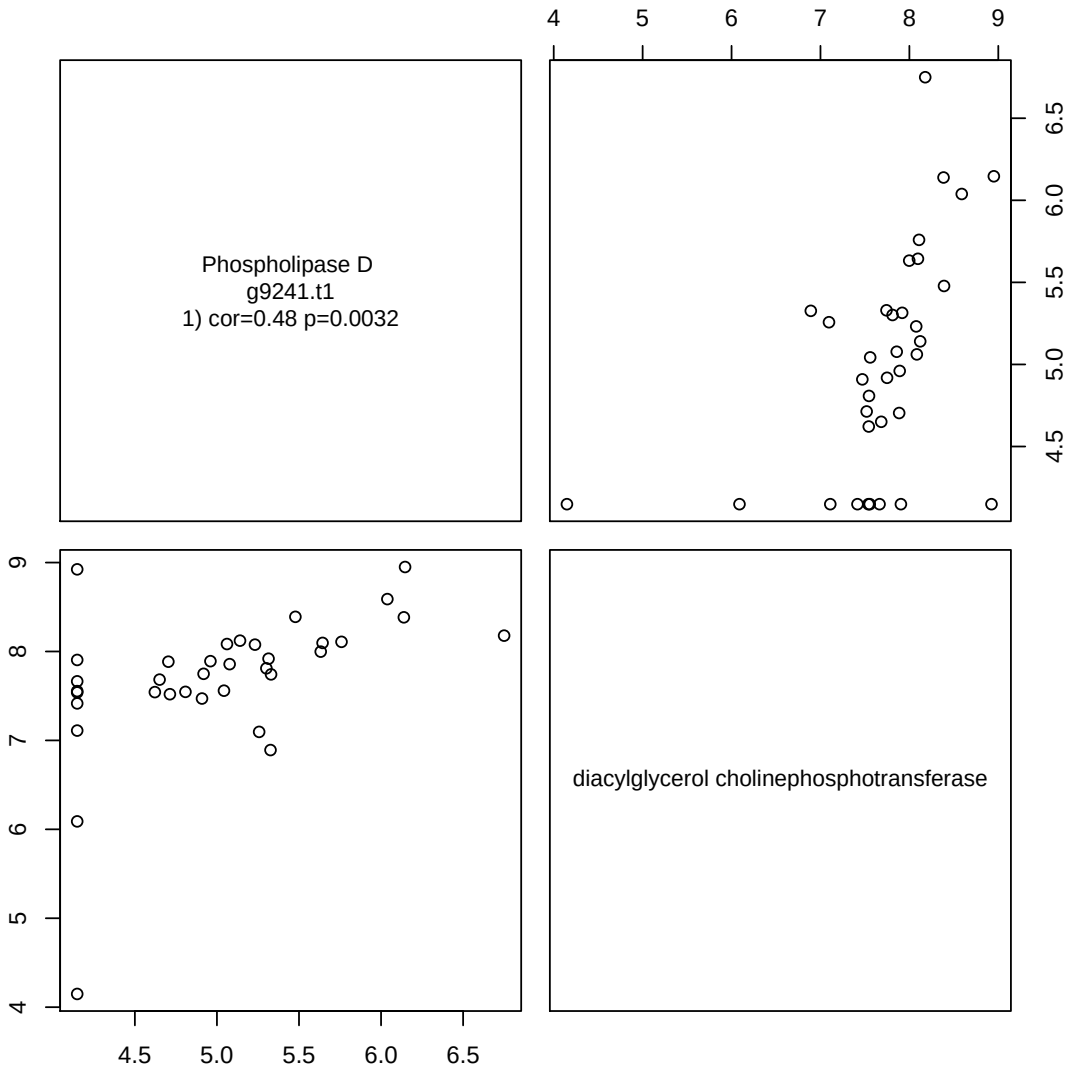


Fig. S13 Pearson correlation between the normalized expressions of the two fungal enzymes, the phospholipase D and the dacylglycerol cholinephosphotransferase degradating the phosphatidylcholine.

Supplementary note S1: Fungal intraradical colonization staining and measurement in *M. esculenta* roots

A sample of approximately 200mg of non-tuberized roots was taken during harvest for each of the 208 plants. These roots were used for the measurement of intraradical fungal colonization. In order to not only confirm the presence of the fungi in each treatments and the absence in the control, but also to measure colonization rate for each isolate. Clean roots of each plants separated in individual eppendorf were soaked in a KOH 10% solution for 4-6 hours in a water bath at 90°C, the KOH solution was changed regularly, until it reached a clear color. After removing completely the KOH solution, HCL 1% was added for 5 min and then replaced with a Trypan blue solution for overnight straining (Trypan blue solution: 333ml of lactic acid, 333ml of ddH_20, 333ml of glycerol and 0.5g of Trypan blue). Trypan blue solution was then replaced by lactic acid 80% for long-term storage. Colonization was measure following an INVAM protocol (International Culture Collection of (Vesicular) Arbuscular Mycorrhizal Fungi; https://invam.wvu.edu/methods/mycorrhizae/mycorrhiza-root-length, last access 02.03.17, Giovannetti and Mosse, 1980). Roots of each plant were randomly spread on petri dish with a grid with squares of 5mm. The numbers of intersection of the roots with the grid were counted with a binocular (brand), as well as the number of roots intersecting the grid that contained a fungal structure (arbuscule, hyphe, spores, vesicule). The number of roots intersecting the grid with a fungal structure divided by the total number of intersection between the roots and the grid multiplied by 100, gave us the colonization rate.

Supplementary note S2: RNA-seq quality control and libraries removal

The quality of the data was inspected in fives ways, in order to assess good quality of the control and the colonized fungal treatments and to remove potentially contaminated control and un-colonized treatments. First of all for samples used for plant morphology as well as for RNA-seq, control were kept only if their colonization assessed by straining was strictly zero, similarly the treatments were keeps only if the colonization by staining was different from zero. Second, colonization was also asses as a percentage of mapped reads to *R. irregularis* compare to the mapped reads of *M. esculenta*. The percentage of *R. irregularis* reads compare to *M. esculenta* was used to discard controls that were considered as contaminated if reaching over 1%. Third, in combination to the reads percentage, the PCA of *R. irregularis* normalized counts including the CTL treatments was used to visualized and discard the potential contaminated control. Fourth, we used FreeBayes (Garrison and Marth, 2012) in order call

SNPs of the *R. irregularis* mapped reads to confirm the presence of the right *R. irregularis* isolate in the right treatment.

Five, we finally searched in the main results 11 well-known plant gene (Gst1, Lec5, Scp1, Pt4, Vapyrin, Flot4, Ann2, Lec7, Glp1, Ram2, Ha1) commonly differently expressed during AMF symbiosis, to confirm the efficiency of the treatments (Hogekamp and Küster, 2013).

Chapter 5

*Ri*MRE: A rare endosymbiotic bacteria in the arbuscular mycorrhiza fungi model: *Rhizophagus irregularis*

Romain Savary[1], Frédéric G. Masclaux[1, 2] and Ian R. Sanders[1].

[1]Department of Ecology and Evolution, Biophore Building, University of Lausanne, 1015 Lausanne, Switzerland. [2] Vital-IT Group, SIB Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland.

Abstract

Arbuscular mycorrhizal fungi (AMF) play an important role in ecosystems by forming an endosymbiotic relationship with most terrestrial plants. This symbiosis is often accompanied by a third partner, an endobacteria inhabiting the fungal cytoplasm. These mollicutes-related endobacteria (MRE) were recently found present in a large number of AMF species, spanning the entire AMF phylogeny and were suggested to have a parasitic nature. It was previously thought that the model AMF species, *Rhizophagus irregularis,* had lost this endobacteria. After screening ddRADseq data obtained from 81 *R. irregularis* isolates and closely related species from across the globe, two isolates, from which one *R. irregularis* isolate, were found to harbour a homogenous haplotype population of MRE. The presence was confirmed by cloning and sequencing of the 16S ribosomal RNA gene. Previously published data on these isolates tended to confirm the hypothesis of a parasitic nature of MRE. Genomes of the *R. irregularis* associated MRE (*Ri*MRE) and its' host, *R. irregularis,* were sequenced to better understand this rare case of the presence of MRE within *R. irregularis,* a species purported to have lost this symbiosis. Genome assembly and analysis are on-going and might reveal interesting features of the intimate life of endosymbionts with reduced genome size and potential horizontal gene transfer (HGT).

Keywords*:* RiMRE, endosymbionte, bacteria, MRE, arbuscular myccorhizal fungi, AMF

Introduction

Arbuscular mycorrhizal fungi (AMF) are well-described symbionts of most plant species, exchanging hard-to-attain soil nutrients for photosynthates. This symbiosis is central to terrestrial ecosystems, as AMF diversity influences the diversity of plant communities, as well as their structure and productivity (van der Heijden et *al.,* 1998). A high variation in AMF phenotypic traits and plant responses were observed (Koch et *al.,* 2006, Klironomos et *al.,* 2003) and could be the result of differences in AMF genetic background (Savary et *al.,* submitted) as well as due to the presence of endobacteria. Indeed, endobacteria are specialized symbionts, present in all AMF clades (Naumann et *al.,* 2010). Moreover, it was shown that the presence of one endobacteria in an AMF isolate, increased fungal fitness (Salvioli et *al.,* 2016) and thereby could impact the outcome of the AM symbiosis.

There has recently been more attention given to these bacteria due to evidence of horizontal gene transfer (HGT) between the AMF genome and the MRE bacterial genomes as a result of their intimate relationship (Naito et *al.,* 2015, Torres-Cortés et *al.,* 2015). This mechanism was proposed as a major force in the evolution of these vertically transmitted bacteria (Torres-Cortés et al., 2015).

Recent studies describe two main species of endobacteria, exhibiting the typical reduced-gene repertoire of a host-depend symbiont (Naito et *al.,* 2015). The first, *Candidatus* Glomeribacter gigasporum (*Ca*Gg) has only been described in the AMF family *Gigasporaceae* and was classified as a member of the *Burkholderia* (Bianciotto et *al.,* 2003, Mondo et *al.,* 2012), a bacterial genus harbouring another fungal endobacteria of *Rhizopus microspores* (Partida-Martinez and Hertweck, 2005). The second, a Mollicute-related endobacteria (MRE), was found to be present in nearly the entire Glomeromycota phylogeny, including *Gigasporaceae* (Naumann et *al.,* 2010, Desirò et *al.,* 2014, Toomer et *al.,* 2015) and are related to a number of animal and plant pathogenic bacteria within the *Mycoplasma* genus (Naumann et *al.,* 2010,). Only few AMF species, such as *Rhizophagus irregularis* were found to be free of these bacteria (Naumann et *al.,* 2010, Toomer et *al.,* 2015).

R. irregularis is a model species of the Glomeromycota, being the first AMF species whose genome was sequenced and publicly available (Tisserant et *al.,* 2013). Furthermore, it is one of the most frequently found species in natural and agricultural soils and one of the easiest to propagate via *in-vitro* culture (Lin *et al.,* 2014). Despite intense research efforts on this species, the presence of *R. irregularis* MRE (*Ri*MRE) within an isolate has not previously been reported. It has been hypothesized that endobacteria played a major role in the evolution

of AMF, as they have been maintained internally and transmitted for more than 100 million years across different AMF lineages (Torres-Cortés et *al.,* 2015). It was therefore suggested that they probably confer an increase of fitness to their host (Naumann et *al.,* 2010), which was recently confirmed for *Ca*Gg, in *G. margarita,* whose presence results in higher sporulation rates, bioenergetic capacity and ATP production (Salvioli et *al.,* 2016). In contrast, Toomer *et al.,* 2015, suggest that the MRE are possible parasites instead of mutualists, as they harbour a high genetic diversity, typical of parasites inherited through vertical transmission who maintain genetic diversity through recombination and horizontal transmission. However, to date no detailed comparison of AMF fitness with regard to MRE presence or absence has been done to confirm a parasitic nature. *R. irregularis* represents an interesting case due to the fact that this endosymbiotic interaction with endobacteria was thought to be lost in this AMF species.

Identifying *R. irregularis* isolates able to form and maintain this symbiosis and quantifying fitness benefits and costs of *Ri*MRE would be a first step in understanding the true nature of the interaction along the symbiosis continuum. Furthermore, studying *Ri*MRE population diversity and sequencing their genomes would also shed light on the degree of HGT that occurred between homologous regions of *R. irregularis* and *Ri*MRE genomes.

To confirm the presence and absence of endobacteria in *R. irregularis* from highly diverse geographical origins, we use ddRAD-seq data on (81) 57 isolates of *R. irregularis* and 22 isolates of related species such as *Rhizophagus intraradices*, *Rhizophagus proliferus* and *Rhizophagus* sp. LPA8-CH3 (Wyss et *al.,* 2016, Savary et *al.,* submitted) and screened for reads mapping to either the MRE genome or the *Ca*Gg genome. Using ddRAD-seq is a non-selective method of sequencing, allowing for sequencing of individuals associated with the organism of interest, such as symbionts or the microbiome. As isolates used for generating ddRAD-seq data originate from sterile, *in vitro* culture, detection of bacterial reads belonging to MRE or *Ca*Gg were easily mapped. Isolates containing traces of endobacteria by mapping were then confirmed by PCR amplification, cloning and Sanger sequencing. These results were discussed in the context of the large literature on *R. irregularis* and on common isolate used in this study. Once confirmed, the genomes of these bacteria were then sequenced to identify potential HGT events. These genomes are being assembled and analysed and will probably resulted in interesting information. Together, this information will allow a better understanding of the endosymbiotic life of these *Ri*MRE bacteria.

Material and Methods

ddRAD-seq and workflow

Recently published ddRAD-seq data on a large population of 81 *R. irregularis* isolates and related species obtained from the bioprojects PRJNA268659 and PRJNA326895 (Wyss et *al.*, 2016, Savary et *al.*, submitted) were used to detect the presence of MRE reads within AMF reads. Read quality check and trimming was done following Savary et *al.*, submitted. The demultiplexed reads of each isolate and each of the three to five replicate were mapped using Novoalign software (Hercus, 2011) to the *Dh*MRE genome, an MRE present in *Dentiscutata heterogama* (Torres-Cortés et al. 2015) and on *Ce*MRE (*Claroideoglomus etunicatum* MRE), *Rc*MRE (*Rhizophagus clarus* MRE) and *Rv*MRE (*Racocetra verrucosa* MRE) genomes (Naito et al. 2015). Reads were also mapped to the other type of AMF-endobacteria found in *Gigaspora margarita*, *Ca*Gg (*Candidatus* Glomeribacter gigasporum) as well to another bacteria also of the genus *Burkholderia* and commonly found in soil (*Burkholderia terrae*, Nazir et *al.*, 2012). SAM files were converted to BAM files and indexed with Samtools (Li *et al.*, 2009). Summary tables generated by Novoalign with the number of mapped reads per replicate to each genome were combined to detect candidate isolates with endobacteria (Table 1). Visualisation of mapped reads was done using IGV (Robinson et al. 2011)

DNA extraction and 16S rRNA and MBLFP gene amplification

Sterile, two-compartment plates were produced for all 81 *R. irregularis* isolates and related species and grown for at least 3 months to obtain sufficient spores for DNA extraction. DNA extraction was performed as described in Savary et *al.*, submitted. The 16S rRNA gene was amplified using primers 109F (109F-1 and 109F-2) and 1184R (1184R-1, 1184R-2 and 1184R-3, Naumann et *al.*, 2010). The MBLFP gene was amplified using MBLFP18f and MBLFP644r (Toomer et *al.*, 2015) primers. A new reverse primer for the 16S rRNA (16SrRNA-RcMRE: 5'-AGTTACCTTGGCAGTCTGC-3') was designed on the basis of the *Rc*MRE 16S rRNA sequence specific to its' genome (Naito et *al.*, 2015). Similarly, a new forward (MBLFP-RcMRE-f, 5'-GAAAYYGGAGAAAAAACTGAYYTAGYYAA-3') and reverse (MBLFP-RcMRE-r, 5'-GARGCATGMARTAWKTCYTCT-3') degenerate primer set for MBLFP were designed based on the homologue MBLFP gene found in *Rc*MRE genome. The following protocol; a PCR mix of 23µl with Qiagen reagents (1X Buffer, 1.5mM $MgCl_2$, 0.2 mM dNTPs, 0.5 µM forward and reverse primer, 0.4 U of *Taq* polymerase) was applied to 2µl of 1 to 25ng/µl of DNA of each of the isolates. The

amplification on a Biometra T1 Thermocycler PCR machine was performed under the following conditions: 5 min at 95°C followed by 35 cycles at 95°C for 10s, 50°C for 30s and 72°C for 2min, a final elongation step at 72°C for 10 min. A slight modification for MBLFP was necessary, by changing the annealing temperature to 49°C and by adding $MgCl_2$ for a final concentration of 2.25 mM. Purification and Sanger sequencing was performed by GATC biotech (Germany). Alignment and sequence cleaning was performed using MEGA 7 (Tamura *et al.*, 2013).

Localisation of the MRE

To confirm the strict endocytoplasmic localisation of MRE, AMF lines carrying MRE, based on ddRAD-seq mapping and Sanger sequencing, were re-extracted from sterile two-compartment plates under a UVC sterilized laminar hood. After dissolution of the fungal compartment for 30 min in citrate buffer (0.0062 M of citric acid anhydrous and 0.0028 M of sodium citrate tribasic dihydrate), the pellet of spores and hyphae were washed with ddH_20, which was used as PCR control. The pellet was surface sterilized with 2% chloramide-T for 3 min, and subsequently washed with 0.03% of Penicillin-streptomycin before being immersed in a 2mL Eppendorf tube filled with ddH_20. Each tube was sonicated for 1 min. The ddH_20 used for sonication served as a second control to detect potential contaminants or external MRE during PCR. Clean spores were finally extracted following the manufacturer's specifications with the DNeasy plant mini kit (QIAGEN).

Cloning and sequencing of the haplotypic population of MRE

MRE were suggested to harbour a large diversity of haplotypes within one AMF isolate (Naito et *al.*, 2015). Thus, cloning of the 16S rRNA amplicon was performed with StrataClone PCR Cloning Kit (Agilent Technologies®) for each isolate carrying an MRE and for which we were able to amplify the 16S region. The cloning protocol was carried out following the manufacturer's specifications.

Phylogeny of 16S rRNA and MBLFP genes

A maximum-likelihood phylogeny was inferred in MEGA7 with a bootstrapping method of 500 replicates (Kumar et *al.*, 2016) using 16S rRNA sequences obtained from Sanger sequencing as well as publicly available sequences (Naumann et *al.*, 2010, Toomer et *al.*, 2015, Naito et *al.*, 2015). A second phylogeny, performed using the same guidelines, was built with MBLFP sequences generated in this study along with previously published

sequences (Toomer et *al.,* 2015) and homologue sequences present within the four different MRE genomes. All trees were edited with Figtree v1.4.2 (Rambaut & Drummond, 2009) and Adobe Illustrator (2014.1.0 Release).

Results

ddRADseq mapping

From the 81 AMF isolates, spanning 4 AMF species, we were able to detect and confirm MRE in 2 isolates, A2 and LPA8 and this with consistency across three to five biological replicates. The highest mean percentage of paired-end reads mapped for both isolates were mapped to the *Rc*MRE genome (Table 1, Fig. 1). One other isolate, CH3 showed signs of MRE reads in all three replicates of ddRAD-seq, but only when the *Rc*MRE genome was used for mapping (Table 1, Fig. 1). Two other isolates (IRTA102, and BEG53) were probably contaminated by exogenous bacterial reads, as an inconsistency among replicates was found in the percentage of read mapping. With some outliers replicates having a higher percentage than other. Across 5 LPA8 replicates, the mean percentage of reads mapping to the 5 different endobacteria genomes was low, with a maximum mean of 0.13% for paired-end reads mapped to *Rc*MRE genome. As expected, we were unable to detect the presence of *Ca*Gg.

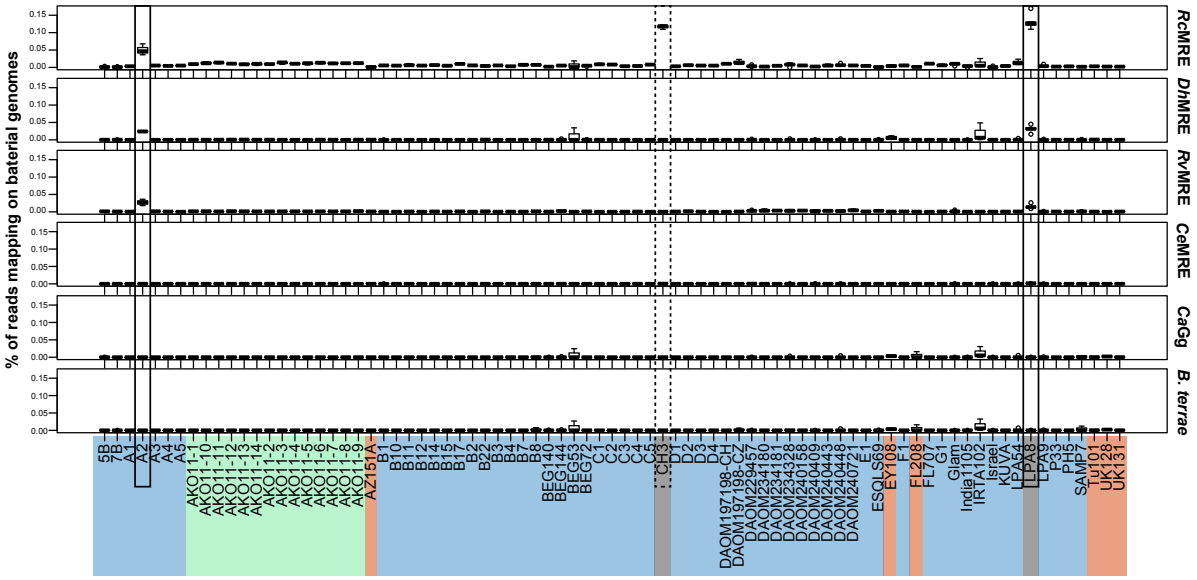


Figure 1. Boxplot of the percentage of ddRAD-seq reads from each of the 81 *R. irregularis* and related species (n=3-5) mapped to four MRE genomes, one *Ca*Gg and one common soil bacterial genome. Percentages are calculated, as the proportion of reads mapped to bacterial genomes compared to the total number of read in a given sample. The colour corresponds to the different species (blue = *R. irregularis*, grey = *R. sp.* LPA8-CH3, red = *R. intraradices* and green = *R. proliferus*).

16S rRNA and MBLFP phylogenies and haplotype populations within isolates

According to both 16S rRNA and MBLFP phylogenic trees, the MRE identified within *R. irregularis,* isolate A2, and within *R. sp.* LPA8-CH3 isolate LPA8, are both closely related to the MRE found in *R. clarus*. Supporting results obtained from mapping ddRAD-seq sequences onto bacterial genomes, we were unable to amplify these two markers in any of the other 79 isolates.

Cloning of A2 and LPA8 16S rRNA amplicon using forward primer 109F-1 and the newly designed reverse primer, 16SrRNA-RcMRE, resulted in homogeneous sequence with only a single MRE haplotype found within A2 (n=8) and LPA8 (n=10).

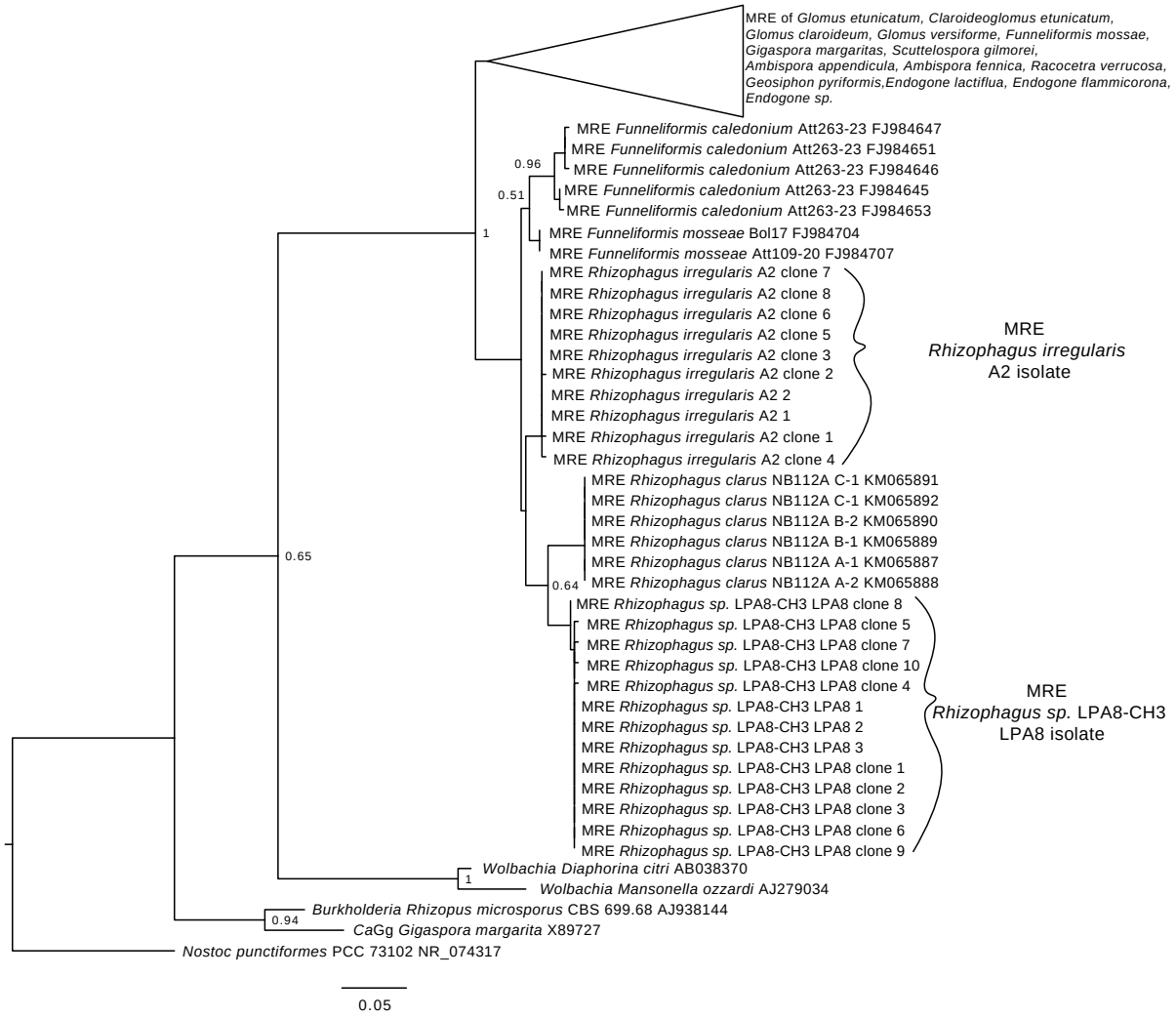


Figure 2. Maximum-likelihood phylogeny of the MRE 16S rRNA sequences rooted with *Nostoc punctiformes*. Cloned sequences of MRE present in *R. irregularis* A2 isolate and *R. sp.* LPA8-CH3 LPA8 cluster with cloned sequences of *R. clarus* obtained from Naito et *al.*, 2015. Here, is also included MRE sequences from Naumann et *al.,* 2010 and Toomer et *al.,* 2015.

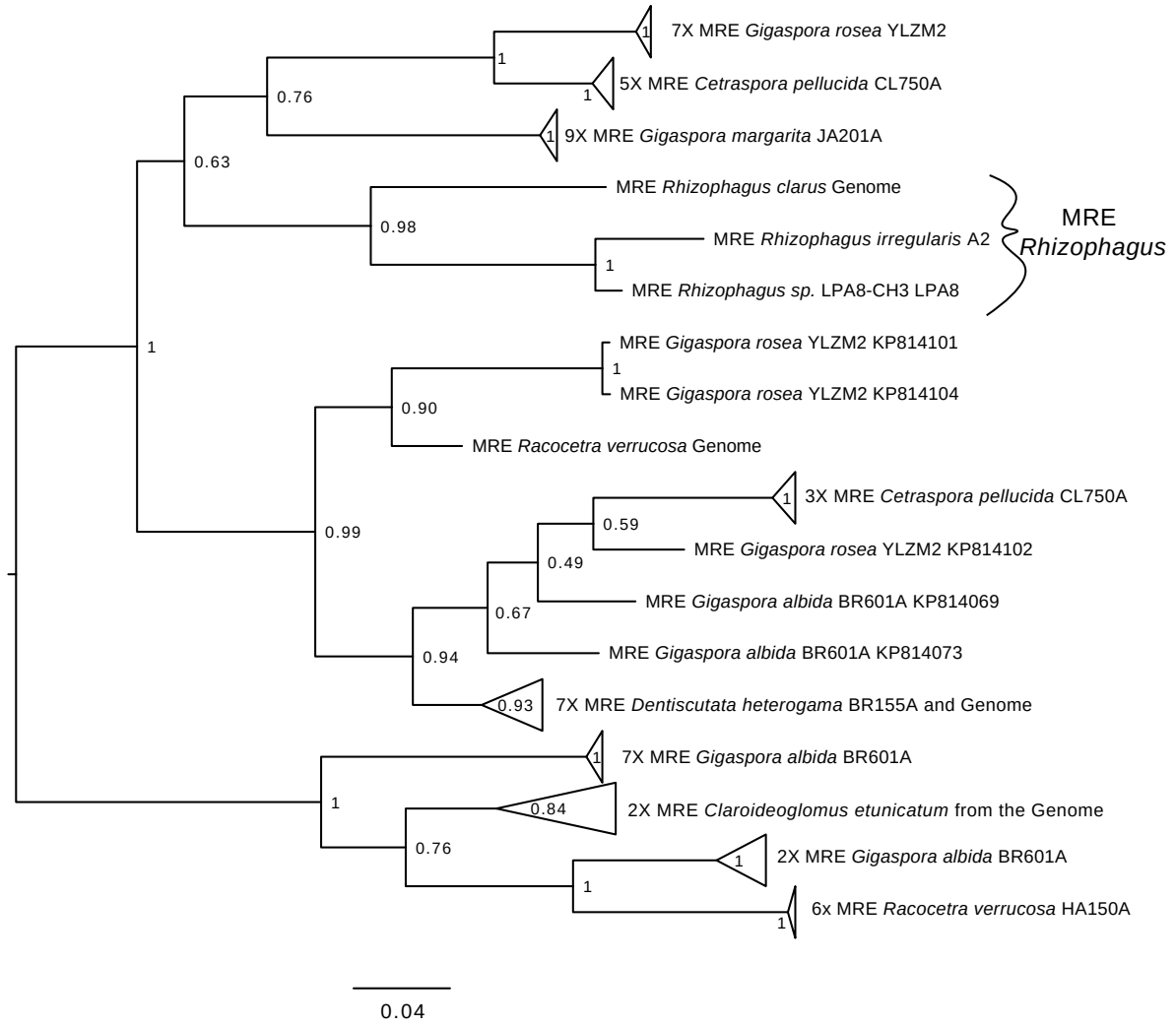


Figure 3. Maximum-likelihood phylogeny of the MRE MBLFP sequences mid-point rooted.
MBLFP sequences obtained from Toomer et *al.,* 2015 are included as well as the MBLFP sequences of four sequenced MRE genomes.

Table 1. Summary table of AMF species, origin of each isolate, mean % of ddRADseq reads proper pair across replicates mapping to 5 bacterial genomes and 16S and MBLFP confirmation. Due to it size the table could be found in electornical format, upon request.

Localisation of MRE bacteria and control for contaminants

As a control, AMF isolates potentially having MRE in their cytoplasm were extracted independently in sterile conditions. Water used for washing and sonication was included as controls during DNA extraction and PCR reactions. No PCR product was detected in this way for the washed water and the sonicated water neither for the 16S rRNA nor for the MBLFP, thus suggesting that no contaminant or MRE are presence outside the AMF host.

Discussion

R. irregularis, despite being a model AMF fungal species, has never been reported to form a symbiosis with an endobacteria. We have shown that this symbiosis is possible however poorly frequent. The A2 isolate that was found to harbour a homogenous haplotypic population of MRE is part of the well-studied *R. irregularis* Swiss population (Croll et *al.,* 2008, 2009, Wyss et *al.,* 2015, Ropars et *al.,* 2016). These Swiss isolates were isolated in 1999 from a single Swiss field-site and have been shown to have high phenotypic and genotypic variability (Koch et *al.,* 2004) as well as they were able to produce a high variability in plant growth responses (Koch et *al.,* 2006). Given the presence of this endobacteria, it may be possible that a part of this variation in plant growth could be influenced by *Ri*MRE endosymbiont.

The results of this study are consistent with a vertical transmission of the *Ri*MRE whose population is maintained within the fungal cytoplasm. This is consistent with supposed co-evolution of the host and the bacterial endosymbiont. Indeed, 16S rRNA sequences of *Ri*MRE present in *R. irregularis* isolate A2, and the candidate species *R. sp.* LPA8-CH8 (Savary et *al.,* submitted) LPA8, reveal a high relatedness to MRE found within other closely related AMF species, *Rhizophagus clarus*, *Funneliformis caledonium* and *Funneliformis mossae*. The relatedness of these MRE strains has been confirmed by analysing the MBLFP gene. Similarly, the highest percentage of paired-end reads mapped from these isolates was onto the *Rc*MRE genome (Naito et *al.,* 2015). A third AMF isolate, CH3, could contain an endosymbiont given the amount of reads mapped to *Rc*MRE. This MRE could possibly be related to the endosymbiont found within *R. clarus,* as reads could not be mapped with other MRE genomes. However, this hypothesis will have to be further investigated since DNA amplification was not possible either with 16S rRNA or MBLFP primer pairs.

The endobacteria are probably found in low quantity within A2, LPA8 and CH3, based on the disproportionate low number of bacterial MRE reads compare to the number of AMF fungal reads.

It has been shown in certain cases that an AMF could host many diverged haplotypes of one MRE within the cytoplasm (Naito et *al.,* 2015). Indeed, by cloning the 16S rRNA gene of MRE living in *C. etunicatum,* Naito et *al.,* 2015 found that this AMF species hosted at least 5 MRE haplotypes, coexisting in the cytoplasm. In contrast, *R. clarus* and *R. verrucosa* both contain a single homogenous population of haplotype endobacteria within the fungal

cytoplasm (Naito et *al.,* 2015). In the case of MRE found within *R. irregularis* and *R. sp.* LPA8-CH3, the population was also shown to be homogenous based on 16S rRNA sequences, as in *R. clarus* as well. It could mean that the population was not enriched by horizontal haplotype transfer as suggested for many other isolates (Toomer et *al.,* 2015).

Toomer et *al.,* 2015, suggested that MRE could be considered as parasite of AMF based on high intra-host diversity, recombination and co-diversification. More recently, another type of endobacteria restricted to the Gigasporaceae family, *Ca*Gg, was shown to increase fitness of its' AMF host (Salvioli et *al.,* 2016), and thus, could be considered as a mutualist of AMF. This observation was only possible thanks to a fungal isolate being cured of the endobacteria. In the field of MRE research, no such system was previously available to evaluate the nature of their symbiosis. However, fungal phenotypic information on many isolates of *R. irregularis* is published and could inform us how this relationship might fit within the spectrum of symbiotic interactions. One striking results of Koch et *al.,* 2004 could be explained by the presence of the endobacteria. Koch et *al.,* measured fungal phenotypic traits on 16 isolates of the Swiss *R. irregularis* population. All these 16 isolates were screened in this study for the presence of MRE. Only a single isolate, A2 among the isolate of Koch et *al.,* was harbouring an MRE. This isolate is the only isolate having strongly reduced spore production compared to other isolates and having a spore production per hypha length ratio that is almost null. Similarly, in a greenhouse experiment (Savary et *al.,* in prep), A2 and 11 other isolates, representing a large phylogenetic diversity of *R. irregularis,* were inoculated on a clonal variety (NGA16) of cassava (*Manihot esculenta*). The colonization rate of this plant was found to differ strongly depending on the different isolate and clades and to be related to the phylogeny. Two phylogenetic clades showed high colonization rates with a mean colonization level of ~50%, with one exception, isolate A2, exhibiting a particularly low colonization rate of ~25%.

The low colonization rate combined to the almost null spore production of A2 suggest that the presence of the endobacteria within this isolate comes at a fitness cost to A2 and tends to support the hypothesis of a parasitic nature of the MRE (Toomer et *al.,* 2015). Moreover MRE are related to the *Mycoplasma* bacterial lineage, a lineage harbouring only parasitic bacteria of plants and animals.

At the mutualistic end of the symbiosis continuum, one of the few other cases of endobacteria in a fungus was found to be related to the mutualist *Ca*Gg. *Burkholderia rhizoxinica* is found in *Rhizopus microspores* and reach even a stronger level of mutualism, where the presence of

the bacteria is obligatory for spore production as well as mycotoxin production (Lackner et *al.*, 2009).

Inspection of geographical distribution of MRE hyplotypes, revealed no genetic differences between continents (Toomer et *al.*, 2015). This is not surprising given recent phylogenetic data from AMF isolates from across the globe, which yielded no sign of endemism or continental structure, but rather suggest that highly similar AMF genotypes can be found in distant locations (Davison et *al.*, 2015 and Savary et *al.*, submitted). For this reason, it is likely the MRE symbiosis is a highly maintained symbiosis, which pre-dates continental separation.

Of the 59 *R. irregularis* isolates, only A2 was found harbouring this symbiont, implying that other *R. irregularis* have lost the symbiont by unknown means. This rarity of endoparasites within this particular AMF clade could have been one evolutionary innovation which contributes to their ecological success compared to other AMF clades still harbouring the endobacteria. Indeed, *R. irregularis* and close relatives are cosmopolitan and often present in large quantities.

In the future, it would be interesting to take advantage of the *in-vitro* status of the A2 isolate to create several lines cured of their endobacteria in order to investigate the role of these bacteria in increasing or decreasing the fitness of the fungus as well as to inspect the presence/absence of this bacteria on plant growth response and plant community structure. It would allow a better understanding of the role of these bacteria in this trinary interaction, involving three Kingdoms. Still, caution should be taken in the method used to cure AMF from their endobacteria. In the past, sub-culturing of single spores from the parent for at least five to six generations was used to cure the *Gigaspora margarita* isolate BEG34 (Lumini et al., 2007), however it has been shown that single spore culturing can result in a genetic modification progeny in *R. irregularis* compare to the parental isolate (Angelard et *al.*, 2010). Moreover, this change in genetic information could impact drastically the host plant response, even if parental isolates of *R. irregularis* already lack endobacteria. For example, these segregated lines, presumed to be clonal, were shown to contain a high diversity of nuclei which likely arose from differences in nucleotype frequencies assembling during spore formation leading to a bottleneck of genetic information in a single spore (Angelard et *al.*, 2010). Thus, the phenotypic comparison between an isolate containing an endobacteria (AMF-B+) and the same isolate cured of endobacteria (AMF-B-) should only be made if a large number of biological replicates of both populations is available, as the variability in the genetic background could over pass the effect of the presence or absence of the endobacteria.

Perhaps a better option, would be to cure the isolate of endobacteria *via* antibiotics, which would not likely influence the segregation of nuclei frequencies.

Acknowledgement

We would like to thank William Robbins for the helpful comments on the manuscript and Jeremy Bonvin for the help in fungal culturing. This study was funded by the Swiss National Science Foundation (Grant number: 31003A_162549 to IRS).

References

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:** 1216-1221.

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:** 121-124.

Croll D, Wille L, Gamper HA, Mathimaran N, Lammers PJ, Corradi N *et al* (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:** 672-687.

Croll D, Giovannetti M, Koch AM, Sbrana C, Ehinger M, Lammers PJ *et al* (2009). Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus Glomus intraradices. *New Phytologist* **181:** 924-937.

Davison J, Moora M, Ouml, pik M, Adholeya A, Ainsaar L *et al* (2015). Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science* **349:** 970-973.

Desiro A, Salvioli A, Ngonkeu EL, Mondo SJ, Epis S, Faccio A *et al* (2014). Detection of a novel intracellular microbiome hosted in arbuscular mycorrhizal fungi. *Isme Journal* **8:** 257-270.

Hercus C. (2011). Novoalign V2.07: http://www.novocraft.com (accessed date June 20 2011)

Klironomos JN (2003). Variation in plant response to native and exotic arbuscular mycorrhizal fungi. *Ecology* **84:** 2292-2301.

Koch AM, Kuhn G, Fontanillas P, Fumagalli L, Goudet J, Sanders IR (2004). High genetic variability and low local diversity in a population of arbuscular mycorrhizal fungi. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 2369-

2374.

Koch AM, Croll D, Sanders IR (2006). Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. *Ecology Letters* **9:** 103-110.

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

Lackner G, Mobius N, Scherlach K, Partida-Martinez LP, Winkler R, Schmitt I *et al* (2009). Global Distribution and Evolution of a Toxinogenic Burkholderia-Rhizopus Symbiosis. *Applied and Environmental Microbiology* **75:** 2982-2986.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:** 2078-2079.

Lin K, Limpens E, Zhang ZH, Ivanov S, Saunders DGO, Mu DS *et al* (2014). Single Nucleus Genome Sequencing Reveals High Similarity among Nuclei of an Endomycorrhizal Fungus. *Plos Genetics* **10**.

Lumini E, Bianciotto V, Jargeat P, Novero M, Salvioli A, Faccio A *et al* (2007). Presymbiotic growth and sporal morphology are affected in the arbuscular mycorrhizal fungus Gigaspora margarita cured of its endobacteria. *Cellular Microbiology* **9:** 1716-1729.

Mondo SJ, Toomer KH, Morton JB, Lekberg Y, Pawlowska TE (2012). Evolutionary stability in a 400-million-year-old heritable facultative mutualism. *Evolution* **66:** 2564-2576.

Naito M, Morton JB, Pawlowska TE (2015). Minimal genomes of mycoplasma-related endobacteria are plastic and contain host-derived genes for sustained life within Glomeromycota. *Proceedings of the National Academy of Sciences of the United States of America* **112:** 7791-7796.

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

Nazir R, Hansen MA, Sorensen S, van Elsas JD (2012). Draft Genome Sequence of the Soil Bacterium Burkholderia terrae Strain BS001, Which Interacts with Fungal Surface Structures. *Journal of Bacteriology* **194:** 4480-4481.

Partida-Martinez LP, Hertweck C (2005). Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* **437:** 884-888.

Rambaut A. & Drummond A. (2009). FigTree v1.3.1.

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G *et al* (2011). Integrative genomics viewer. *Nature Biotechnology* **29:** 24-26.

Ropars J, Toro KS, Noel J, Pelin A, Charron P, Farinelli L *et al* (2016). Evidence for the sexual origin of heterokaryosis in arbuscular mycorrhizal fungi. *Nature Microbiology* **1**.

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

Savary R, Masclaux FG, Wyss T, Droh G, Cruz Corella J, Machado AP, Morton JB Sanders IR. Submitted**.** A population genomics approach shows widespread geographical distribution of cryptic genomic forms of the symbiotic fungus *Rhizophagus irregularis***.** submitted to *Isme journal* 2017.

Savary R, Dupuis C., Masclaux FG, Mateus Gonzalez I. Rojas Tayo EC. Sanders IR. In prep. Dual RNA-seq of Cassava (*Manihot esculenta*) roots colonized by the *Rhizophagus irregularis* genetic diversity. In prep.

Tisserant E, Malbreil M, Kuo A, Kohler A, Symeonidi A, Balestrini R *et al* (2013). Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* **110:** 20117-20122.

Toomer KH, Chen XH, Naito M, Mondo SJ, den Bakker HC, VanKuren NW *et al* (2015). Molecular evolution patterns reveal life history features of mycoplasma-related endobacteria associated with arbuscular mycorrhizal fungi. *Molecular Ecology* **24:** 3485-3500.

Torres-Cortes G, Ghignone S, Bonfante P, Schussler A (2015). Mosaic genome of endobacteria in arbuscular mycorrhizal fungi: Transkingdom gene transfer in an ancient mycoplasma-fungus association. *Proceedings of the National Academy of Sciences of the United States of America* **112:** 7785-7790.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T *et al* (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396:** 69-72.

Wyss T, Masclaux FG, Rosikiewicz P, Pagni M, Sanders IR (2016). Population genomics reveals that within-fungus polymorphism is common and maintained in populations of the mycorrhizal fungus Rhizophagus irregularis. *Isme Journal* **10:** 2514-2526.

Chapter 6

General discussion

Findings summary

Up to now, the diversity of AMF at large scales suggested a low global endemism. However these results were controversial due to the low resolution of the markers used. In chapter 2, I found, using high-resolution markers, that two AMF species, *R. intraradices* and *R. irregularis* could be found on different continents and at highly distant locations. Moreover *R. irregularis* showed similar patterns of low endemism at the within species level. My results confirmed that at the species, and even at the intra-species level, long distance dispersion and low endemism is possible in AMF. With this I, thus, identified a previously undescribed within *R. irregularis* diversity under the form of four *R. irregularis* genetics groups, this intra-specific diversity was suggested to be ecologically functional and relevant.

In chapter 3, I found that the within AMF species diversity identified in chapter 2 resulted in a gradual change in plant community response observed along the *R. irregularis* phylogeny in mesocosms simulating European calcareous grassland. This effect was stronger for some isolates, especially in the repression of dominance of one plant species, leading to a greater evenness in the community. These results support the view that AMF intra-species diversity is ecologically relevant.

In the chapter 4, I found that a general transcriptomic reprogramming occurred in the roots of the crop plant *M. esculenta* in response to genetically different *R. irregularis* forms. This reprogramming took the form of a change in plant transcription of genes such as proteases that could control the progression of the fungi, receptors that could detect the fungi, increase in the fatty acid synthesis and in transporters to exchange nutrients such as ammonium and potentially phospholipids. Despite these clear signs of molecular changes, the response was not visible at the plant phenotypic level, for simple traits. In the fungi, sharp differences in transcriptomic strategies were observed across the diversity of *R. irregularis*, suggesting different molecular strategies to invade the host.

Finally in chapter 5, I took advantage of ddRAD-seq data produced in chapter 2, to search for the presence of endobacteria that were previously shown to impact AMF fitness. Mollicute-related endobacteria (MRE) were found in most AMF clades but were found lacking in *Rhizophagus irregularis* (Naumann et *al.*, 2010, Toomer et *al.*, 2015). I found the presence of an MRE in only one isolate of the *R. irregularis* species. This bacteria that I refer to RiMRE, presented a homogenous population with a single haplotype. In other AMF, MRE were suggested as being potential parasites. Thus, *R. irregularis* mostly successfully eradicated it

potential parasite, and this single case of preservation of the symbiosis probably at the limit of extinction of this bacterium could teach us a lot about maintenance of endosymbiotic life.

Rhizophagus irregularis genomic, phenotypic and transcriptomic diversity

Rhizophagus irregularis is the model AMF fungus and its genome and transcriptome were the first to be sequenced and assembled. However, given the importance of the symbiosis and the large body of literature using *R. irregularis* to understand the AMF symbiosis it is surprising that almost no information is available on the natural genomic diversity of this species. One population of *R. irregularis* in Switzerland has been studied (Koch et *al.,* 2004, 2006, Croll et *al.,* 2008, Roger et *al.,* 2013). Those studies showed, at the small-scale, an unexpected high level of genetic diversity and a high variability in plants growth responses to inoculation with these fungi. Moreover, one isolate genotyped in that study originated from Canada. Despite it origins, it displayed a higher relatedness to certain isolates in the Swiss population than some Swiss isolates to each other. This single observation raised the question about the endemism of AMF. Could the same species of AMF disperse far and have survived in different climates, soils and with completely different plant hosts? A first positive clue toward this came from a global scale AMF community profiling study, showing a surprisingly low endemism for 93% of the species (Davison et *al.,* 2015). However, this last study combined with the findings of other studies (Munkvold et *al.,* 2004, Koch et *al.,* 2006, Mensah et *al.,* 2015) raised a new question. Do we use a sufficiently fine molecular resolution to define biologically and ecologically relevant species? Indeed the study of Davison et *al.,* (2015) was criticized, as the low-resolution marker they used would not be sufficient to define the relevant level of the species. It was even suggested by Bruns and Taylor in 2016 that “all of these “species” (morphologicaly or single-marker defined) in the Glomeromycota are actually collections of fairly distantly related taxa”. In chapter 2, I provide evidence that for this species it is the case. Indeed, I have shown that within *R. irregularis* species clear genomic divergence is found, separating this species into four main genetic groups. I have shown also that this intra-specific divergence, as well as the divergence between the species studied (*R. irregularis, R. intraradices and R. proliferus*) would not be detected with the marker and the clustering methods traditionally used by ecologists to make AMF community profiling such as used in the study by Davison et *al.,* 2015. However, even though I found more genetic diversity that previously recorded in the *Rhizophagus* genus, and within *Rhizophagus irregularis* I did not

find (Bruns & Taylor, 2016) “that endemic species will be discovered to be the rule”. In contrast, I also found (Davison et *al.,* 2015) low endemism in *R. irregularis* and *R. intraradices* as well as in the genetic groups of *R. irregularis.*

The phenotypic measurements also showed significant differences in extra-radical mycelial (ERM) density among the different genetic groups of *R. irregularis*. This is important for species that rely on the trade of the nutrients collected by ERM. The nutrients taken up by the ERM are traded in the plant roots. I also uncovered differences in colonization rates levels among genetic groups of *R. irregularis* that were tightly linked to their phylogeny. Some clades were better colonizers than others on *M. esculenta* and on *Festuca pratensis*.
The genomic, ERM and colonization differences among *R. irregularis* genetic groups combined with their clear differences in their whole transcriptomes are strong clues to suggest that these different genetic groups are evolutionarily stable units and could be considered as different species. I propose according to this findings that AMF which were previously considered as a poor groups of fungi in species number summing only 300-1600 species (van der Heijden et *al.,* 2015) could be composed of a way larger number of species. This would make sense from a biological and ecological perspective. These results also bring important clues to suggest that level of relevant biological and ecological diversity is not considered in ecological studies.

Rhizophagus irregularis relatedness effect on plant communities

Strong effects of AMF diversity on plant diversity, community productivity and structure were shown, demonstrating the potentially important role of AMF in ecological processes. However most ecological studies used available isolates from morphologically described species (Maherali & Klironomos, 2007) or sometimes from unidentified species (van der Heijden et *al.,* 1998). Such AMF diversity might not be the most relevant or not the only important AMF level of diversity. In chapter 3, I found that by using intra-specific diversity of different related isolates of *R. irregularis,* a signal within the phylogeny can be found within a plant community. This effect largely targets the dominant plants of the community, with a phylosignal of colonization and with an overall negative responsiveness that can be strong or less strong depending on the *R. irregularis* genetic groups. This negative responsiveness impacted the community by increasing or decreasing community evenness. In

other words, closely related isolates within an AMF species affect a whole plant community in the same way, with some clade of *R. irregularis* resulting in an higher or lower plant community evenness. I also found that increasing the phylogenetic diversity of *R. irregularis* isolates within a community increased the community evenness until a certain point. Then at high phylogenetic diversity the evenness decreased. These findings point towards the importance of within-AMF species diversity impacting plant-plant interactions in plant communities.

Rhizophagus irregularis and *Manihot esculenta* diversity of molecular tools to perform symbiosis.

In the last years, more and more plant genes have been found that are involved in the AM symbiosis. Some of these are necessary to allow establishment of symbiosis (Gutjahr et *al.,* 2012) and many are conserved across all mycorrhizal plants (Delaux et *al.,* 2014, Bravo et *al.,* 2016) and differentially express (Hogekamp & Küster, 2013). The knowledge of the "gene machinery" necessary to make the symbiosis is well established. However, on the fungal side most of the machinery is still unknown. In chapter 4, I have tried to fill this gap by using the variability of "colonizer type" (mean colonization rate of an isolate) and genetic groups found within *R. irregularis*.

I found a particularly interesting gene that is differentially transcribed among some *R. irregularis* genetic groups. That is co-expressed in a gene module linked to AMF colonizer type. This gene is a homologue of one of the 64 effectors found in *R. clarus* (Sędzielewska & Brachmann, 2016). This secreted protein; a chitin deacytelase was shown in another fungus (Cord-Landwehr et *al.,* 2016) to help the fungi to modify chitin oligomers released by plant chitinases. By doing this, the fungus will not be recognized by the plant receptors, *de facto* inactivating the plant defences. The variability in the expression of this gene has probably played an important role in the ability of AMF to invade different host plants, Cord-Landwehr et *al.,* even generalized this to most plant-invading fungi.

Another interesting point raised by the dual RNA-seq of *M. esculenta* and *R. irregularis*, is the ability of the plant to specially increase the transcription of genes involved in fatty acid synthesis during the symbiosis. Indeed for a long time AMF were thought to receive

carbohydrate in the form of sugar from their host plants. However, the combined sequencing of different AMF genomes recently revealed that type 1 fatty acid synthase (FAS) is absent in *R. irregularis* (Wewer et *al.,* 2014), thus, leading to the hypothesis that the fungi might be unable to correctly perform the synthesis of fatty acid and could obtain the fatty acid from the plant. This hypothesis is supported by the fact that plant lipid content was observed to change drastically during the AM symbiosis and that a number of fatty acid genes are conserved across mycorrhizal plants (Bravo et *al.,* 2016). In 2017, Bravo et *al.,* (2017) suggested that conserved plant fatty acid synthesis genes (FatM and RAM2) and transport (STR and STR2) were involved in the synthesis and the transfer of a monoacylglycerol (MAG) to the fungi.

In my experiment, all the *M. esculenta* genes implicated in fatty acid synthesis were up regulated when the plant was inoculated with any genetic form of the *R. irregularis* species. The path suggested by Bravo et *al.,* (2017) for the transfer of MAG was similarly up regulated in *M. esculenta* and tended to support the view of the transfer of MAG to the fungi. However, I found that another path was possible for the transfer of lipids, by the intermediate of a gene with the similar functionality of RAM2, but this time leading to the synthesis of phospholipids (PL) and to phosphatidylcholine (PC). Under the form of lyso-phosphatidylcholine, PC was shown to be used in AMF as a signal (Drissner et *al.,* 2007), thus suggesting that it could be transferred from the fungi to the plant and inversely. Moreover PC production was shown to be increase in inoculated roots (Wewer et *al.,* 2014). Despite that not all the genes in the path of PC synthesis were found to be up-regulated, keys gene of the synthesis, transfer and degradation from the plant to the fungi were found to be differentially expressed (DE) between colonized and non-colonized plants and among *R. irregularis* genetic groups. First, a plant phospholipid transporting ATPase was found to be DE and correlated to the expression of two fungal phospholipid transporting ATPases. This suggests a synchrony in the transport of PL. Second, in the fungi, different enzymes of the phosphatidylcholine degradation and transformation were found to be differentially expressed among *R. irregularis* genetic groups. These enzymes, lead either to the transformation of PC into diacylglycerol (DAG) and then to triacylglycerol (TAG), the main lipid found in AMF ERM, or lead the product of decomposition of the PC to the glycolysis.

The hypothesis that PC is transported to the fungi is supported by the observation that an increased production of PC occurs when the plant is inoculated with AMF, and that the main phospholipid in the ERM is PC (Wewer et *al.,* 2014). In chapter 4, I proposed that PC used as a signal molecule in AMF symbiosis could be recycled by the fungi for the formation of DAG and TAG or used in glycolysis. However, this use could vary from one genetic group to

another, with a stronger or weaker ability of transporting and transforming the phospholipids. I also found that genetic groups with a lower ability to transform the PC could use another enzyme; a glycerol kinase, to obtain carbon from the glycerol. Such a hypothesis is supported by the paper of Wei et *al.*, (2004) who showed the ability of a semi-biotrophic fungus to obtain glycerol from the plant and survive only by this uptake.

Rhizophagus irregularis and its rare endosymbiotic bacteria; *Ri*MRE

It is known that highly specialized and obligate endobacteria occur in most AMF. Two types of endosymbiotic bacteria occur in AMF. The first, the *Ca*Gg, was shown to be restricted to the Gigasporaceae family. The second, the mollicute related endobacteria (MRE, Naumann et al., 2010), are found in almost all the species of AMF including Gigasporaceae. *Ca*Gg were suggested to be mutualists of the Gigasporaceae, increasing the fitness of their host (Salvioli et *al.*, 2016) while MRE were considered as parasites of AMF due to their high within population haplotypic diversity (Toomer et *al.*, 2015). Despite the large distribution of MRE in almost all AMF, *R. irregularis* is reported as having lost its endosymbiont. This potentially suggests a fungal success in the host-parasite arms race. The presence of such potentially parasitic or mutualistic bacteria could potentially perturb the symbiosis and the plant response to these fungi.

In order to unravel the presence of these endobacteria in chapter 5, I used ddRAD-seq data generated in chapter 2 and standard 16S primers. The presence of the bacteria was confirmed with reads of the ddRAD-seq, 16S sequences and another nuclear gene in one *R. irregularis* isolate among the large-scale *R. irregularis* population. An additional MRE was found in one isolate of *Rhizophagus* sp. LPA8-CH3. A second isolate of the *Rhizophagus* sp. LPA8-CH3 also showed signs of the presence of MRE. However their presence was not confirmed neither with 16S gene or the other nuclear gene. The *R. irregularis* isolate inhabited by the endobacteria was shown to exhibit highly reduce spore production (Koch et *al.*, 2004) and colonization rate (chapter 4) compared to related isolates. This tends to confirm that MRE could have a parasitic life style (Toomer et *al.*, 2015). However, the 16S sequences of the two confirmed isolates were cloned and resulted in a homogenous population of haplotypes, more specifically characterizing a non-essential mutualistic partner or maybe simply representing a relic of their parasitic ancestry.

Future work and conclusions

In this work I suggested that the AMF diversity evaluated to 300-1600 species could be underestimated. In order to confirm my prediction, it would be important to undertake other studies looking at the genomic diversity within other AMF species. As up to now the study of the genomic of AMF needed a large amount of DNA, it would be crucial to isolate more species *in vitro* to obtain enough DNA. Many isolates are already available in different AMF bank and small enterprises and could be shared among researchers. Another option would be to develop new methods to perform ddRAD-seq or whole genome sequencing, on single spores directly extracted and cleaned from the soil, to avoid a potential selection of fungi during the isolation process. Such a protocol could potentially use whole genome amplification (WGA). The amplification and the sequencing of the genome of single nuclei (Lin et *al.*, 2014) was already achieved, suggesting that such protocol could be developed, thus, avoiding the time consuming and possibly biased work of single spore isolation.
In my work I have confirmed, at the within-species level, that there is weak or no endemism in *R. irregularis* as suggested from the global AMF community profiling study of Davison et *al.*, (2015). However, it would be important to isolate or obtain isolates of a single species from uncultivated soils around the world in order to perform deep molecular characterization and to confirm my results and those of Davison et *al.*, (2015).

In chapter 3, I found that plant community structure could be impacted by intra-specific diversity of one AMF species. However, it would be essential to investigate the effect of intra-specific diversity on plant community biomass production and diversity. In that sense, an open field plot should be used in order to avoid a resource limitation found in pot experiments. Similar experiments as performed by van der Heijden et *al.*, 1998 but with a within species diversity and a know relatedness of the isolates could bring good insight of ecological role of within-AMF species diversity.
The findings of chapter 4 highlighted new insights into how the AMF-plant symbiosis works and showed a certain variation in the molecular strategies used by the different genetic groups of *R. irregularis*. I also proposed many genes that could be essential for the nutrient acquisition by the fungus as well as a gene that could enable the fungi to hide from the plant defence. It would be important to further investigate the role of these potentially important genes. Moreover I found modules of fungal and plant genes that interact, it would be

interesting to see how much of the module structure is maintained in other similar experiments, in order to confirm the central fungal and plant genes involved in the symbiosis.

In the last chapter, I highlight the presence of an endobacteria within one isolate of *R. irregularis*. This rare case of endobacteria within *R. irregularis*; a species that is thought to have eradicated its endosymbiotic bacteria, could highlight the nature of the endosymbiotic life and how one AMF species finally succeeded in eradicating a bacterial partner with which it had co-evolved for the last 400Myrs. Moreover, the sequencing of other genomes of MRE suggested that some homologue genes of *R. irregularis* were horizontally transmitted to the MRE (Naito et *al.*, 2015, Torres-Cortes et *al.*, 2015). By sequencing the genome of the RiMRE and its fungal host I should be able to answer important questions of how and why the symbiosis has evolved and was maintained. Trans-kingdom horizontal gene transfers (HGT) that were observed between AMF and their endosymbiotic bacteria were found by comparing the bacterial gene repertories to the only fungus sequenced to date, *R. irregularis*, but never to their own AMF host. It would, thus, be important to sequence the *R. irregularis* host genome and its endosymbiotic partner genome.

Having an *in vitro* isolate retaining an endobacteria is a unique chance to study the impact of the endobacteria on fungal host fitness as well as to study the results of this interaction on the plant. Indeed, it was shown experimentally that the type of AMF endobacteria restricted to Gigasporaceae had the ability to increase AMF host fitness. Although the MRE contained in the *R. irregularis* isolate and in most AMF species, was suggested to be parasitic, this fact was never demonstrated. By removing the bacteria with antibiotics or by segregation over generations, it could be possible to study the effect of the bacteria on the phenotype of the same isolate with or without the bacteria. It would be also possible to see the resulting effect on plants, by producing large quantities of *in vitro* inoculum with and without the bacteria. Further studies could even sequence the transciptome of the fungus with bacteria or without bacteria in order to unravel which gene are expressed while the bacteria is present. This would give more clues to understand the nature of this intriguing and certainly important driver of the evolution of AMF.

General introduction and discussion references

Al-Karaki G, McMichael B, Zak J (2004). Field response of wheat to arbuscular mycorrhizal fungi and drought stress. *Mycorrhiza* **14:** 263-269.

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:** 1216-1221.

Bago B, Pfeffer PE, Abubaker J, Jun J, Allen JW, Brouillette J *et al* (2003). Carbon export from arbuscular mycorrhizal roots involves the translocation of carbohydrate as well as lipid. *Plant Physiology* **131:** 1496-1507.

Bravo, A. Brands M. Wewer V. Dörmann P. Harrison MJ. 2017. Arbuscular mycorrhiza-specific enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular mycorrhiza. *New Phytologist*. doi:10.1111/nph.14533

Bravo A, York T, Pumplin N, Mueller LA, Harrison MJ (2016). Genes conserved for arbuscular mycorrhizal symbiosis identified through phylogenomics. *Nature Plants* **2**.

Bruns TD, Taylor JW (2016). FUNGAL SYMBIONTS Comment on "Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism". *Science* **351**.

Ceballos I, Ruiz M, Fernandez C, Pena R, Rodriguez A, Sanders IR (2013). The In Vitro Mass-Produced Model Mycorrhizal Fungus, Rhizophagus irregularis, Significantly Increases Yields of the Globally Important Food Security Crop Cassava. *Plos One* **8**.

Cord-Landwehr S, Melcher RLJ, Kolkenbrock S, Moerschbacher BM (2016). A chitin deacetylase from the endophytic fungus Pestalotiopsis sp efficiently inactivates the elicitor activity of chitin oligomers in rice cells. *Scientific Reports* **6**.

Croll D, Wille L, Gamper HA, Mathimaran N, Lammers PJ, Corradi N *et al* (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:** 672-687.

Croll D, Giovannetti M, Koch AM, Sbrana C, Ehinger M, Lammers PJ *et al* (2009). Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus Glomus intraradices. *New Phytologist* **181:** 924-937.

Croll D, Sanders IR (2009). Recombination in Glomus intraradices, a supposed ancient asexual arbuscular mycorrhizal fungus. *Bmc Evolutionary Biology* **9**.

Davison J, Moora M, Ouml, pik M, Adholeya A, Ainsaar L *et al* (2015). Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science* **349:** 970-973.

Delaux PM, Varala K, Edger PP, Coruzzi GM, Pires JC, Ane JM (2014). Comparative Phylogenomics Uncovers the Impact of Symbiotic Associations on Host Genome Evolution. *Plos Genetics* **10**.

Drissner D, Kunze G, Callewaert N, Gehrig P, Tamasloukht M, Boller T *et al* (2007). Lyso-phosphatidylcholine is a signal in the arbuscular mycorrhizal symbiosis. *Science* **318:** 265-

268.

Gonzalez-Chavez C, Harris PJ, Dodd J, Meharg AA (2002). Arbuscular mycorrhizal fungi confer enhanced arsenate resistance on Holcus lanatus. *New Phytologist* **155:** 163-171.

Govindarajulu M, Pfeffer PE, Jin HR, Abubaker J, Douds DD, Allen JW *et al* (2005). Nitrogen transfer in the arbuscular mycorrhizal symbiosis. *Nature* **435:** 819-823.

Gutjahr C, Radovanovic D, Geoffroy J, Zhang Q, Siegler H, Chiapello M *et al* (2012). The half-size ABC transporters STR1 and STR2 are indispensable for mycorrhizal arbuscule formation in rice. *Plant Journal* **69:** 906-920.

Hajiboland R, Aliasgharzadeh N, Laiegh SF, Poschenrieder C (2010). Colonization with arbuscular mycorrhizal fungi improves salinity tolerance of tomato (Solanum lycopersicum L.) plants. *Plant and Soil* **331:** 313-327.

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB (2001). Molecular evidence for the early colonization of land by fungi and plants. *Science* **293:** 1129-1133.

Hogekamp C, Kuster H (2013). A roadmap of cell-type specific gene expression during sequential stages of the arbuscular mycorrhiza symbiosis. *Bmc Genomics* **14**.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ *et al* (2006). Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443:** 818-822.

Karandashov V, Bucher M (2005). Symbiotic phosphate transport in arbuscular mycorrhizas. *Trends in Plant Science* **10:** 22-29.

Kiers ET, Duhamel M, Beesetty Y, Mensah JA, Franken O, Verbruggen E *et al* (2011). Reciprocal Rewards Stabilize Cooperation in the Mycorrhizal Symbiosis. *Science* **333:** 880-882.

Klironomos JN, McCune J, Hart M, Neville J (2000). The influence of arbuscular mycorrhizae on the relationship between plant diversity and productivity. *Ecology Letters* **3:** 137-141.

Klironomos JN (2003). Variation in plant response to native and exotic arbuscular mycorrhizal fungi. *Ecology* **84:** 2292-2301.

Koch AM, Kuhn G, Fontanillas P, Fumagalli L, Goudet J, Sanders IR (2004). High genetic variability and low local diversity in a population of arbuscular mycorrhizal fungi. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 2369-2374.

Koch AM, Croll D, Sanders IR (2006). Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. *Ecology Letters* **9:** 103-110.

Koricheva J, Gange AC, Jones T (2009). Effects of mycorrhizal fungi on insect herbivores: a meta-analysis. *Ecology* **90:** 2088-2097.

Kruger M, Kruger C, Walker C, Stockinger H, Schussler A (2012). Phylogenetic reference data for systematics and phylotaxonomy of arbuscular mycorrhizal fungi from phylum to species level. *New Phytologist* **193:** 970-984.

Lin K, Limpens E, Zhang ZH, Ivanov S, Saunders DGO, Mu DS *et al* (2014). Single Nucleus Genome Sequencing Reveals High Similarity among Nuclei of an Endomycorrhizal Fungus. *Plos Genetics* **10**.

Maherali H, Klironomos JN (2007). Influence of Phylogeny on fungal community assembly and ecosystem functioning. *Science* **316:** 1746-1748.

Mariotte P, Vandenberghe C, Kardol P, Hagedorn F, Buttler A (2013). Subordinate plant species enhance community resistance against drought in semi-natural grasslands. *Journal of Ecology* **101:** 763-773.

Mensah JA, Koch AM, Antunes PM, Kiers ET, Hart M, Bucking H (2015). High functional diversity within species of arbuscular mycorrhizal fungi is associated with differences in phosphate and nitrogen uptake and fungal phosphate metabolism. *Mycorrhiza* **25:** 533-546.

Munkvold L, Kjoller R, Vestberg M, Rosendahl S, Jakobsen I (2004). High functional diversity within species of arbuscular mycorrhizal fungi. *New Phytologist* **164:** 357-364.

Naito M, Morton JB, Pawlowska TE (2015). Minimal genomes of mycoplasma-related endobacteria are plastic and contain host-derived genes for sustained life within Glomeromycota. *Proceedings of the National Academy of Sciences of the United States of America* **112:** 7791-7796.

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

O'Malley MA (2015). Endosymbiosis and its implications for evolutionary theory. *Proceedings of the National Academy of Sciences of the United States of America* **112:** 10270-10277.

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

Parniske M (2008). Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nature Reviews Microbiology* **6:** 763-775.

Pozo MJ, Azcon-Aguilar C (2007). Unraveling mycorrhiza-induced resistance. *Current Opinion in Plant Biology* **10:** 393-398.

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

Rodriguez A, Sanders IR (2015). The role of community and population ecology in applying

mycorrhizal fungi for improved food security. *Isme Journal* **9:** 1053-1061.

Roger A, Colard A, Angelard C, Sanders IR (2013). Relatedness among arbuscular mycorrhizal fungi drives plant growth and intraspecific fungal coexistence. *Isme Journal* **7:** 2137-2146.

Ropars J, Toro KS, Noel J, Pelin A, Charron P, Farinelli L *et al* (2016). Evidence for the sexual origin of heterokaryosis in arbuscular mycorrhizal fungi. *Nature Microbiology* **1**.

Rosendahl S (2008). Communities, populations and individuals of arbuscular mycorrhizal fungi. *New Phytologist* **178:** 253-266.

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

Sanders IR, Rodriguez A (2016). Aligning molecular studies of mycorrhizal fungal diversity with ecologically important levels of diversity in ecosystems. *Isme Journal* **10:** 2780-2786.

Song FQ, Li JZ, Fan XX, Zhang Q, Chang W, Yang FS *et al* (2016). Transcriptome analysis of Glomus mosseae/Medicago sativa mycorrhiza on atrazine stress. *Scientific Reports* **6**.

Tisserant E, Malbreil M, Kuo A, Kohler A, Symeonidi A, Balestrini R *et al* (2013). Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* **110:** 20117-20122.

Toomer KH, Chen XH, Naito M, Mondo SJ, den Bakker HC, VanKuren NW *et al* (2015). Molecular evolution patterns reveal life history features of mycoplasma-related endobacteria associated with arbuscular mycorrhizal fungi. *Molecular Ecology* **24:** 3485-3500.

Toro KS, Brachmann A (2016). The effector candidate repertoire of the arbuscular mycorrhizal fungus Rhizophagus clarus. *Bmc Genomics* **17**.

Torres-Cortes G, Ghignone S, Bonfante P, Schussler A (2015). Mosaic genome of endobacteria in arbuscular mycorrhizal fungi: Transkingdom gene transfer in an ancient mycoplasma-fungus association. *Proceedings of the National Academy of Sciences of the United States of America* **112:** 7785-7790.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T *et al* (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396:** 69-72.

van der Heijden MGA, Martin FM, Selosse MA, Sanders IR (2015). Mycorrhizal ecology and evolution: the past, the present, and the future. *New Phytologist* **205:** 1406-1423.

Vandenkoornhuyse P, Leyval C, Bonnin I (2001). High genetic diversity in arbuscular mycorrhizal fungi: evidence for recombination events. *Heredity* **87:** 243-253.

Vandenkoornhuyse P, Husband R, Daniell TJ, Watson IJ, Duck JM, Fitter AH *et al* (2002).

Arbuscular mycorrhizal community composition associated with two plant species in a grassland ecosystem. *Molecular Ecology* **11:** 1555-1564.

Vandenkoornhuyse P, Ridgway KP, Watson IJ, Fitter AH, Young JPW (2003). Co-existing grass species have distinctive arbuscular mycorrhizal communities. *Molecular Ecology* **12:** 3085-3095.

Wagg C, Jansa J, Stadler M, Schmid B, van der Heijden MGA (2011). Mycorrhizal fungal identity and diversity relaxes plant-plant competition. *Ecology* **92:** 1303-1313.

Wei YD, Shen WY, Dauk M, Wang F, Selvaraj G, Zou JT (2004). Targeted gene disruption of glycerol-3-phosphate dehydrogenase in Colletotrichum gloeosporioides reveals evidence that glycerol is a significant transferred nutrient from host plant to fungal pathogen. *Journal of Biological Chemistry* **279:** 429-435.

Wewer V, Brands M, Dormann P (2014). Fatty acid synthesis and lipid metabolism in the obligate biotrophic fungus Rhizophagus irregularis during mycorrhization of Lotus japonicus. *Plant Journal* **79:** 398-412.

Wyss T, Masclaux FG, Rosikiewicz P, Pagni M, Sanders IR (2016). Population genomics reveals that within-fungus polymorphism is common and maintained in populations of the mycorrhizal fungus Rhizophagus irregularis. *Isme Journal* **10:** 2514-2526.

Curriculum Vitae

Romain Savary
Neuengasse 1, 3011 Bern, Switzerland
romain.savary@unil.ch (+41) 79 398 81 01
Swiss 17.03.88

EDUCATION

12-17 **PhD,** Genomics, transcriptomics and ecology of fungal-plant symbiosis
Dept. of Ecology and Evolution, University of Lausanne, Switzerland.

09-11 **MSc.** in Behaviour, Evolution and Conservation, Avian pollutant and Fish genetics
Dept. of Ecology and Evolution, University of Lausanne, Switzerland.

06-09 **BSc.** in Biology, *University of Lausanne, Switzerland.*

PROFESSIONAL EXPERIENCE

12-17 **PhD** student with Prof. **Ian R. Sanders**:
"*Large-scale genetic diversity of the symbiotic fungus:* Rhizophagus irregularis *and its effect on a plant host and a plant community*"

- Managing project from greenhouse to lab. and to computer analysis
- Writing and publishing projects
- Supervision of Master student thesis
- Master and Bachelor teaching.

11-12 **Laboratory and field assistant, (**Prof. **Nicolas Perrin),** Amphibian genetics

- Fieldwork and supervision of internship student.
- Labwork: PCR, setting up microsatellite protocols, cloning, Sanger sequencing, sequence editing and analysis, microsatellites genotyping and analysis. Supervision of PhD and Master student for labwork.

09-11 **MSc.** Student with Dr. **Lucas Fumagalli** and Dr. **Pierre Bize**

- Master Project (published): "*Stocking activities for the Arctic charr in Lake Geneva: Genetic effects in space and time*"
- First step research project: "*Effect of persistent organic pollutants exposure on Alpine swift (*Apus melba*) reproductive success in two Swiss colonies*"

09- **Infield Guide and Scientific advisor for the Swiss Cetacean Society (SCS)**
Scientific advisor, fieldwork guide and skipper assistant on sailboat for field data collection. Teaching and supervision of up to 10 eco-volunteers for scientific data collection on cetacean and marine wildlife in the Mediterranean Sea.

09 **Internship in KORA** (Coordinated research projects for the conservation and management of carnivores in Switzerland)**,** used of GIS tools for the Lynx, Bern

SKILLS

Scientific and project management skills:

- Expertise in Genetics and next-generation sequencing for Genomics and Transcriptomics.
- Expertise and knowledge in the following taxonomic groups, Amphibians, Fish, Snakes, Plants, Cetacean, Shark and marine wildlife with a will to increase the knowledge to other groups.
- Ability to manage multiple projects at the same time, supervise people in a project, work with a large diversity of researcher and present in front of an audience the results of my work.

Fieldwork

-Field campaign in different country collecting DNA samples of fish, various amphibians or other organisms.
-Field campaign on boat for collecting data on cetacean in the Mediterranean Sea with the Swiss Cetacean Society (SCS) or collecting data on whale shark in the Maldives with the Maldives Whale Shark Research Programme (MWSRP).
-Greenhouse experiments with large number of tropical (Cassava) and local plants.
-Amphibians captivity care taking and breeding

Laboratory

Fungal *in vitro* culture, Plant *in vitro* propagation, DNA & RNA extraction, PCR protocol development and improvement, cloning, microsatellites genotyping, Sanger sequencing, ddRADseq & RNA-seq libraries preparation.

Informatics/ Bioinformatics

-Word, Excel, PowerPoint, Illustrator, Lightroom
-R and Unix languages and use of Vital-IT cluster (https://www.vital-it.ch)
-Sanger sequences analysis from sequence editing to trees
-Microsatellites genotypes reading to population genetic tools
-Command line software for Genomics and Transcriptomics from reads cleaning, mapping, SNPs calling to more advance analytical tools, differential expression analysis.

Writing

Attended to a writing course during my first year of PhD, publications of my Master project and my first chapter of thesis.

Teaching

-During my PhD I have supervised a Master student (1 year) and a large number of biology student (internship of several months).
-I have taught Botany and Mycology and amphibian ecology to Bachelor students for 4 years.
-I have taught for 4 years marine shoreline ecology, a 2-weeks course in Roscoff and molecular biology for 1 year to post-graduate students.
-I have taught marine ecology and cetology to adult during Mediterranean sailing campaigns of one week each.
-I have participated to different exhibitions, vulgarizing science for the public, about symbiosis or marine life.

Languages

French: Mother tongue
English: Working language
German: Intermediate
Italian: Basic understanding

Other	-Driving licence **-**Certificate of Swiss reptile course, KARCH (Coordinated Swiss center of protection and conservation of amphibian and reptiles) **-**Organisation of the Conservation day 2016, a day at the university of Lausanne presenting the different organs of the wildlife conservation **-**Diving PADI open water certificate
Interests/	Large interest in fauna ecology, genetics and conservation with a particular interest for marine fauna
Hobbies	Practicing outdoor sport such as running, snowboarding, surfing, diving, sailing, cycling, swimming and climbing. I like as well travelling, drawing, photography and the music.

PUBLICATIONS

** Les auteurs on contribué à part égal* *IF = Impact factor*

In preparation & submitted

10. **Savary R.*,** Masclaux FG.*, Sanders IR. *Ri*MRE : The genome of the rare *Rhizphagus irregularis* endosymbiotic bateria. (**In prep.** intended for a submission to PNAS**)**

9. **Savary R.,** Dupuis C., Masclaux FG, Mateus Gonzalez I., Rojas Tayo EC., Sanders IR. Effects of intra-specific variation in *Rhizophagus irregularis* on the transcriptomes of mycorrhizas and Cassava in symbiosis. (**In prep.** intended for a submission to Current Biology)

8. **Savary, R.***, Villard, L.*, Sanders IR. Within-species phylogenetic relatedness of a common mycorrhizal fungus predicts plant community structure. (**submitted** to the Journal of Ecology) *IF* = 5.4

Accepted & Published

7. **Savary, R.,** Masclaux, F. G., Wyss, T., Droh, G., Cruz Corella, J., Machado**,** AP., Morton, JB., Ian R. Sanders. A population genomics approach shows widespread geographical distribution of cryptic genomic forms of the symbiotic fungus *Rhizophagus irregularis*. (accepted 2017**, The ISME journal**). *IF* = 9.7

6. **Savary, R.,** Dufresnes, C., Champigneulle, A., Caudron, A., Dubey, S., Perrin, N., Fumagalli, L. Stocking activities for the Arctic charr in Lake Geneva: Genetic effects in space and time. 2017. **Ecology and Evolution**, 7, 5201-5211. *IF* = 2.4

5. Dufresnes C., Majtyka T., Baird S.J., Gerchen J.F., Borzée A., **Savary R.**, Ogielska M., Perrin N., Stöck M. Empirical evidence for large X-effects in animals with undifferentiated sex chromosomes.
2016. **Scientific Reports**, 6 p. 21029. *IF* = 4.3

4. Dufresnes C., Bertholet Y., Wassef J., Ghali K., **Savary R.**, Pasteur B., Brelsford A., Rozenblut-Kościsty B., Ogielska M., Stöck M. et al., Sex-chromosome differentiation parallels postglacial range expansion in European tree frogs (*Hyla arborea*). 2014, **Evolution**, 68 (12) pp. 3445-3456. Peer-reviewed. *IF* = 4.2

3. Brelsford A., Stöck M., Betto-Colliard C., Dubey S., Dufresnes C., Jourdan-Pineau H., Rodrigues N., **Savary R.**, Sermier R., Perrin N. Homologous sex chromosomes in three deeply divergent anuran species. 2013. **Evolution**, 67 (8) pp. 2434-2440. Peer-reviewed. *IF* = 4.2

2. Stöck M., **Savary R.**, Zaborowska A., Górecki G., Brelsford A., Rozenblut-Kościsty B., Ogielska M., Perrin N. Maintenance of ancestral sex chromosomes in Palearctic tree frogs: direct evidence from *Hyla orientalis*. 2013. **Sexual Development**, 7 (5) pp. 261-266. Peer-reviewed. *IF* = 2.0

1. Stöck M., **Savary R.**, Betto-Colliard C., Biollay S., Jourdan-Pineau H., Perrin N, Low rates of X-Y recombination, not turnovers, account for homomorphic sex chromosomes in several diploid species of Palearctic green toads (*Bufo viridis* subgroup). 2013. **Journal of Evolutionary Biology**, 26 (3) pp. 674-682. Peer-reviewed. *IF* = 2.8

CONFERENCES Talks and posters

31st Annual Conference, European Cetacean Society, 2017, Denmark
Swiss Mycorrhizal Symposium (SMS) 2016, Zürich, Oral
Symposium of Ecology and Evolution Doctoral Students (SeeDS), 3MT, Lausanne, Oral
European Society for Evolutionary Biology (ESEB) congress, Lausanne 2015, Poster
international Molecular Mycorrhiza Meeting (iMMM) Cambridge 2015, Poster
33rd New Phytologist Symposium 2015, Zürich, Poster
Population genomics doctoral program meeting 2014, Charmey, Oral
Conference at the Annual meeting of water 2012, Lausanne, Oral

REFERENCES

Prof Ian R. Sanders
Dept. of Ecology and Evolution, University of Lausanne
ian.sanders@unil.ch
+41(0)21 692 42 61

Prof Nicolas Perrin
Dept. of Ecology and Evolution, University of Lausanne
nicolas.perrin@unil.ch
+41(0)21 692 41 84

Dr. Lucas Fumagalli
Dept. of Ecology and Evolution, University of Lausanne
lucas.fumagalli@unil.ch
+41(0)21 692 41 72