Université de Montréal

Comparative Mitochondrial Genomics Toward Understanding Genetics and Evolution of Arbuscular Mycorrhizal Fungi

By

Maryam Nadimi

Département de sciences biologiques
Faculté des Arts et des sciences

Thèse présentée à la Faculté des arts et des sciences
en vue de l'obtention du grade de docteur
en sciences biologiques

November, 2014

Résumé

Les champignons mycorhiziens arbusculaires (CMA) sont très répandus dans le sol où ils forment des associations symbiotiques avec la majorité des plantes appelées mycorhizes arbusculaires. Le développement des CMA dépend fortement de la plante hôte, de telle sorte qu'ils ne peuvent vivre à l'état saprotrophique, par conséquent ils sont considérés comme des biotrophes obligatoires. Les CMA forment une lignée évolutive basale des champignons et ils appartiennent au phylum Glomeromycota. Leurs mycélia sont formés d'un réseau d'hyphes cénocytiques dans lesquelles les noyaux et les organites cellulaires peuvent se déplacer librement d'un compartiment à l'autre. Les CMA permettent à la plante hôte de bénéficier d'une meilleure nutrition minérale, grâce au réseau d'hyphes extraradiculaires, qui s'étend au-delà de la zone du sol explorée par les racines. Ces hyphes possèdent une grande capacité d'absorption d'éléments nutritifs qui vont être transportés par ceux-ci jusqu'aux racines. De ce fait, les CMA améliorent la croissance des plantes tout en les protégeant des stresses biotiques et abiotiques. Malgré l'importance des CMA, leurs génétique et évolution demeurent peu connues. Leurs études sont ardues à cause de leur mode de vie qui empêche leur culture en absence des plantes hôtes. En plus leur diversité génétique intra-isolat des génomes nucléaires, complique d'avantage ces études, en particulier le développement des marqueurs moléculaires pour des études biologiques, écologiques ainsi que les fonctions des CMA. C'est pour ces raisons que les génomes mitochondriaux offrent des opportunités et alternatives intéressantes pour étudier les CMA. En effet, les génomes mitochondriaux (mt) publiés à date, ne montrent pas de polymorphismes génétique intra-isolats. Cependant, des exceptions peuvent exister. Pour aller de l'avant avec la génomique mitochondriale, nous avons besoin de générer beaucoup de données de séquençages de l'ADN mitochondrial (ADNmt) afin d'étudier les méchanismes évolutifs, la génétique des

population, l'écologie des communautés et la fonction des CMA. Dans ce contexte, l'objectif de mon projet de doctorat consiste à: 1) étudier l'évolution des génomes mt en utilisant l'approche de la génomique comparative au niveau des espèces proches, des isolats ainsi que des espèces phylogénétiquement éloignées chez les CMA; 2) étudier l'hérédité génétique des génomes mt au sein des isolats de l'espèce modèle *Rhizophagus irregularis* par le biais des anastomoses ; 3) étudier l'organisation des ADNmt et les gènes mt pour le développement des marqueurs moléculaires pour des études phylogénétiques.

Nous avons utilisé l'approche dite '*whole genome shotgun*' en pyroséquençage 454 et Illumina HiSeq pour séquencer plusieurs taxons de CMA sélectionnés selon leur importance et leur disponibilité. Les assemblages *de novo*, le séquençage conventionnel Sanger, l'annotation et la génomique comparative ont été réalisés pour caractériser des ADNmt complets. Nous avons découvert plusieurs mécanismes évolutifs intéressant chez l'espèce *Gigaspora rosea* dans laquelle le génome mt est complètement remanié en comparaison avec *Rhizophagus irregularis* isolat DAOM 197198. En plus nous avons mis en évidence que deux gènes *cox1* et *rns* sont fragmentés en deux morceaux. Nous avons démontré que les ARN transcrits les deux fragments de *cox1* se relient entre eux par épissage en trans '*Trans-splicing*' à l'aide de l'ARN du gene *nad5* I3 qui met ensemble les deux ARN *cox1.1* et *cox1.2* en formant un ARN complet et fonctionnel. Nous avons aussi trouvé une organisation de l'ADNmt très particulière chez l'espèce *Rhizophagus* sp. Isolat DAOM 213198 dont le génome mt est constitué par deux chromosomes circulaires. En plus nous avons trouvé une quantité considérable des séquences apparentées aux plasmides '*plasmid-related sequences*' chez les Glomeraceae par rapport aux Gigasporaceae, contribuant ainsi à une évolution rapide des ADNmt chez les Glomeromycota. Nous avons aussi séquencé plusieurs isolats de l'espèces *R. irregularis* et *Rhizophagus* sp. pour décortiquer leur position phylogénéque et inférer des relations évolutives entre celles-ci. La comparaison

génomique mt nous montré l'existence de plusieurs éléments mobiles comme : des cadres de lecture '*open reading frames* (mORFs)', des séquences courtes inversées '*short inverted repeats* (SIRs)', et des séquences apparentées aux plasimdes '*plasmid-related sequences* (*dpo*)' qui impactent l'ordre des gènes mt et permettent le remaniement chromosomiques des ADNmt. Tous ces divers mécanismes évolutifs observés au niveau des isolats, nous permettent de développer des marqueurs moléculaires spécifiques à chaque isolat ou espèce de CMA.

Les données générées dans mon projet de doctorat ont permis d'avancer les connaissances fondamentales des génomes mitochondriaux non seulement chez les Glomeromycètes, mais aussi de chez le règne des Fungi et les eucaryotes en général. Les trousses moléculaires développées dans ce projet peuvent servir à des études de la génétique des populations, des échanges génétiques et l'écologie des CMA ce qui va contribuer à la compréhension du rôle primorial des CMA en agriculture et environnement.

Mots clés : champignons mycorhiziens arbusculaires (CMA), génome mitochondrial, mitogénomique comparative, marqueur moléculaire, homoplasmie, anastomose, analyses phylogénétique, remaniement génomique, élément génétique mobile, recombinaison.

Abstract

Arbuscular mycorrhizal fungi (AMF) are the most widespread eukaryotic symbionts that form mutualistic association with majority of plant-roots known as Arbuscular Mycorrhizae. AMF are obligate biotrophs belonging to an ancient fungal lineage of phylum Glomeromycota. Their mycelia are formed by a complex network made by coenocytic hyphae, where nuclei and cell organelles can freely move from a compartment to another. AMF are commonly acknowledged to improve plant growth by enhancing mineral nutrient uptake, in particular phosphate and nitrate and they confer tolerance to abiotic and biotic stresses for plants. Despite their significant roles in ecosystems, their genetics and evolution are not well understood. Studying AMF is challenging due to their obligate biotrophy, their slow growth and limited morphological treats. In addition, intra-isolate genetic polymorphism of nuclear DNA brings another level of complexity for investigation biology, ecology and function of AMF. Genetic polymorphism of nuclear DNA within a single isolate limits the development of efficient molecular markers mainly at lower taxonomic level (i.e. inter-isolate level). Thus, mitochondrial (mt) genomics have been used as an attractive alternative to study AMF. mt genomes have been shown to be homogeneous or much less polymorphic than nuclear DNA in AMF. However, we need to generate large mt sequence datasets in order to investigate its efficiency and usefulness in developing molecular marker toolkits in order to study its dynamic and evolutionary mechanisms as well as population genetics, community ecology and functions of Glomeromycota. In line with these challenges, the objectives of my Ph.D. project were therefore: 1) To investigate mitochondrial genome evolution using comparative mitogenomic analyses of closely related species and isolates as well as phylogenetically distant taxa of AMF; 2) To explore mt genomes inheritance among compatible isolates of the model AMF *Rhizophagus irregularis* through

anastomosis formation; and 3) To assess mtDNA and mt genes for marker development and phylogenetic analyses.

We used whole genome shotgun, 454 pyrosequencing and HiSeq Illimina to sequence some selected AMF taxa according to their importance and availability in our lab collections. *De novo* assemblies, Sanger sequencing, annotation and comparative genomics were then performed to characterize completed mtDNAs. We discovered interesting evolutionary mechanisms in *Gigaspora rosea* in which we found that its mtDNA revealed the fully reshuffled genome synteny compared to *Rhizaphagus irregularis* DAOM 197198 and presence of two fragmented *cox1* and *rns* genes. We demonstrated that two *cox1* transcripts are joined by trans-splicing. We also reported an unusual mtDNA organization in *Rhizophagus* sp. DAOM 213198 whose mt genome consisted of two circular mtDNAs. In addition, we observed a considerable amount of mt plasmid-related sequences in Glomeraceae compared with Gigasporaceae contributing to fast evolution of mtDNA in Glomeromycota. We also sequenced other isolates of *R. irregularis* and *Rhizophagus* sp. in order to unravel their evolutionary relationship and to develop molecular toolkits for their discrimination. Comparative mitogenomic analyses of these mtDNAs revealed the occurrence of many mobile elements such as mobile open reading frames (mORFs), short inverted repeats (SIRs), plasmid-related sequences (*dpo*) that impact mt genome synteny and mtDNA alteration. All together, these evolutionary mechanisms among closely related AMF isolates give us clues to design reliable an efficient intra- and inter-specific markers that are to discriminate closely related AMF taxa and isolates.

Data generated in my Ph.D. project advanced our knowledge on mitochondrial genomes evolution not only in Glomeromycota, but also in Fungal kingdom and in Eukaryotes in general. Molecular toolkits developed in this project will offer new opportunities to study population genetics, genetic exchanges and ecology of AMF, which will contribute to understand the role of

these fungi in nature with potential application in agriculture and environment protection.

Key words: arbuscular mycorrhizal fungi (AMF), mitochondrial genome, comparative mitogenomics, molecular marker, homoplasmy, anastomosis, phylogenetic analyses, genome rearrangement, mobile genetic elements, recombination.

Table of Contents

List of tables

List of figures

List of abbreviations

~: approximately
AM: arbuscular mycorrhiza
AMF: arbuscular mycorrhizal fungi
RNA: acide ribonucléique
mRNA: ARN messager
ATP: adenosine triphosphate
atp: ATP synthetase
BiP: binding protein
BLAST: basic local alignmentBLAST: basic local alignment tool
Bp: base pair(s)
cDNA: complementary deoxyribonucleic acid
cob: apocytochrome b
cox: cytochrome oxidase
C*-terminal: carboxy-terminal
DAOM: agriculture and agri-Food Canada national mycological herbarium
DGGE: denaturing gradient gel electrophoresis
DNA: deoxyribonucleic acid
dNTP: deoxyribonucleoside triphosphate
dpo: DNA polymerase
e.g.: Latin: exempli gratia (English: for example)
HMM: Hidden Markov Model
i.e.: Latin: *id est* (English: that is)
in vitro: Latin: in glass
in vivo: Latin: within the living organism
ITS: internal transcribed spacer
Kb: kilo base pair(s)
LSU: large subunit
µl: micro liter
µM: micro molar
mM: mili molar
ML: Maximum likelihood
mRNA: messenger RNA
mt: mitochondrial
MM: minimal medium
MUCL: mycothèque de l'universite catholique de Louvain
nad: NADH dehydrogenase
NADH: nicotinamide adenine dinucleotide ng: nano gram

NGS: next generation sequencing
NCBI: the national center for biotechnology information
N*-terminal: amino-terminal
NSERC: natural sciences and engineering research council (Canada)
Numts: nuclear mitochondrial DNA
Orf: open reading frames
PCR: polymerase chain reaction
PCG: protein coding gene
RAPD: Randomly Amplified Polymorphic
rDNA: ribosomal deoxyribonucleic acid
RFLP: restriction fragment length polymorphism
Ri T-DNA: transfer DNA transmission from Agrobacterium root-inducing plasmid into the genome of *Dacus carota*
rnl: large ribosomal RNA
rns: small ribosomal RNA
rpm: revolutions per minute
rpo: RNA polymerase gene
rps: ribosomal protein subunit
Rpo: RNA polymerase
rRNA: ribosomal ribonucleic acid
RNA: ribonucleic acid
ROC: root organ culture
RT-PCR: real-time polymerase chain reaction
SSU: small subunit
Spp.: species
Taq: *Thermus aquaticus*
TGGE: temperature gradient gel electrophoresis
Tm: melting temperature
tRNA: transfer ribonucleic acid
TRFLP: terminal restriction fragment length polymorphism
UV: ultraviolet
v/v: volume/volume
w/v: weight/volume
WGA: whole genome amplification

This thesis is dedicated to my mother and father, who had the patience to be far from me to complete this research in past five years. Thank you and I love you.

Acknowledgments

“And still, after all this time, the Sun has never said to the Earth,
"You owe me."
Look what happens with love like that.
It lights up the sky.”
— Hafez

It is my pleasure to thank the many people who made this research possible, first and foremost my supervisor Dr. Mohamed Hijri. I wish to express my most sincere appreciation to his nice attitude and kindness. His smiley face always full me up with energy for continueing my research. I was not only learned scientific knowledge from him but also how patience and calmness can be as effective as anything else. I would like to express very special gratitude to Dr. Franz Lang for transferring his worthy knowledge to me in bioinformatic analyses. Wiothout his assistance I would not be able to persue my research in this field. My appreciation also goes to my husband, Amir Yadghar, my beloved parents, Hamid and Ensieh, my sister, Fatima and my dear parents-in-law, Ali and Elaheh for their love, understanding, patience, and calmness that are the most effective support.

I wish to express my sincere appreciation to knowledgeable postdoctoral fellows Dr. Franck Stefani and Cristina Micali for their support and valuable advices. The work could also not have been completed without the aid of my colleague Dr. Eva Boon, Dr. Denis Beaudet, Dr. Sebastien Halary, Dr. Ivan de la Providencia, Laurence Daubois, Alice Roy-Bolduc, Guillaume Bourdel, dear Stephanie Berthiaume, Iffis Bachir, and Rim Ben Haj Sassi. I would like to thank Dr. Firas Bou Daher for his special supports. I would also acknowledge Dr. Anne Bruneau and Dr. Simon Joly and Dr. Sebastien Renaut and Edeline Gagnon for their assistance in phylogenetic analyses and Dr. Marc St. Arnaud and Dr. Mario Cappadocia for their beneficial advises in my predoctoral

exam. I wish to express my last, but not least, wholehearted thanks to my lovely friends for their motivation, encouragement and support. Finally, I am the most grateful person in the world by having the sweetest son, Ali, during my Ph.D. I am sorry that I did not have enough time for him and I wish he would excuse me.

Chapter 1 - General introduction

1.1 Mycorrhiza

Exploring the plants' rhizosphere reveals that it acts such as a factory and microorganisms and plants activities produce the outcome of this natural factory which is essential to soil, plant and ecosystem health. ‘Mycorrhiza' (Greek word means ‘fungus-root') used for the first time by Franck (1885) to describe the symbiotic association between fungi and root of trees. This association was dated back 400 million years ago based on the fossils records (Remy et al. 1994). Mycorrhizas are the most important and widespread symbiosis between plants and fungi on earth. This mutualistic association provides the fungus with carbohydrates, in a form that they can easily translocate and absorb. In return, fungal mycelia extend out of the roots into the soil where they uptake mineral nutrients and water that will be delivered to plants' root. As a consequence, mycorrhizal fungi play a crucial role in plant nutrient uptake, water relations, ecosystem establishment, plant productivity and diversity. Up to date, seven major important groups of mycorrhizae have been characterized and described, (Peterson et al. 2004) as; (1) arbuscular mycorrhiza, (2) ectomycorrhiza, (3) ectendomycorrhiza, (4) arbutoid mycorrhiza, (5) monotropoid mycorrhiza, (6) ericoid mycorrhiza and (7) orchid mycorrhiza. Different groups of mycorrhizas not only diverge in structural features such as intracellular structure (coil in Arbuscular, ericoid and arbutoid mycorrhizas; peloton in orchid mycorrhiza, fungal peg in monotropoid mycorrhiza and arbuscule in arbuscular mycorrhiza) but also in their preferred host plants and ecosystems which outcome their distribution (Smith and Read 2008). Arbuscular

mycorrhiza is the most important group of mycorrhizae, which is widespread that is found in all ecosystems and concerns more than 80% of plants. Arbuscular mycorrhizal fungi (AMF), is an early divergent fungal lineage that can form mycorrhiza with most important crops in which they improve growth and productivity. Therefore, they were largely used as an alternative to hazardous mechanisms of agriculture and chemical fertilizers (biofertilizers), substituting the development and maintenance of healthy ecosystems. Nowadays, many commercial inoculants of AMF with a wide range of formulations are available in the market.

1.2 Arbuscular mycorrhizal fungi

Arbuscular mycorrhizal fungi (AMF) are a group of high potential microorganisms for agriculture that belong to the ancient phylum Glomeromycota (Schüβler et al. 2001). They are considered as living fossils since they are found to be 460 million years old (Redecker et al. 2000). AMF contributed directly to the evolution and survival of plant species, to the expansion of biodiversity in the earth and consequently to the equilibrium of ecosystems. These microorganisms are ubiquitous putative mutualistic fungi that can form symbiotic associations with the majority of vascular plants (Parniske 2008). However, non-mutualistic or exploitive (i.e. an association in which only the plant profits from the nutrient exchange, (Taylor and Bruns 1999) is a challenging term for mycorrhizal fungi in association with myco-heterotrophic plants that are completely supported by a fungus (Imhof 1999 and Brundrett 2004). The mycelium of these coenocytic (i.e. hyphae that are lacking septa) root symbionts acts as root extension for plants improving the use of soil water and soil minerals from a much larger volume of soil. Improvement in nutrient uptake enhances plant growth and health (Smith 1997) lead to resistance

and controlling the environmental stresses (e.g. drought, salinity and pollution) and plant pathogens (Barea et al. 2002). Moreover, AMF stimulate the fitness of plants in polluted environments (Hildebrandt et al. 1999) and affect soil structure by improving soil quality (Barea et al. 2002). AMF also raises resistance to biotic and abiotic stresses, representing a relevant alternative for sustainable agriculture (Subramanian and Charest 1999, Aliasgharzad et al. 2006). These fungi are declared to be obligate biotrophs, which means they are not able to complete their life cycle and cultivate without a host plant. However, Hildebrandt and colleagues questioned this putative virtue of AMF by finding *G. intraradices* (*G. irregulare*) capable of completing its life cycle in the presence of *Paenibacillus validus* in the absence of a plant root (Hildebrandt et al. 2006). AMF grow a hyphal network both within the root cortex (linear and coiling; (Smith 1997)) and the adjacent soil. The physical properties of the mycelium (a high surface-to-volume ratio), enhance uptake of nutrients more than the plant's root system by exploring a larger volume of soil (Bolan 1991, Tuomi et al. 2001). AMF were previously considered as asexual (clonal) organisms (Gandolfi et al. 2003). However, increasing number of evidences such as detection of genetic recombination, the presence of ‘core meiotic genes' and retrotransposons (Halary et al. 2013), questioned this phenomenon. These fungi form spores containing hundreds of nuclei that are most likely haploid (i.e., one set of chromosome in a single nucleus) (Hijri and Sanders 2004). Analyses of Kuhn et al. (2001) by fluorescence *in situ* hybridization (FISH) revealed two highly variable variants of internal transcribed spacer 2 (ITS2) in different nuclei of *Scutellospora castanea*, suggesting AMF heterokaryosis. Pawlowska and Taylor (2004) believed this fungus to be homokaryotic because of the analysis of *POL*-like sequences from *Glomus etunicatum*, which show that all sequence variants were present in offsprings. However, Sanders and Croll (2010) explained the results of Pawlowska and Taylor (2004) in a review, and concluded that AMF are most likely heterokaryotic. Hijri & Sanders

(2005) and Boon et al. (2015) support the heterokaryosis hypothesis in AMF by reanalyzing POL-like genetic sequence variations in *G. etunicatum* (and *R. irregularis* DAOM 197198). It is important to note that AMF do not seem to go through a bottleneck of genetic variation (Marleau et al. 2011); in other words, no observations of a single nucleus stage have been reported in AMF life history. The lack of a genetic bottleneck is considered as a potential source for AMF extreme intra-isolate genetic diversity (Kuhn et al. 2001). However, Lin et al. (2015) sequenced the genomes of individual nuclei of *R. irregularis* DAOM 197198 and they found a low level of polymorphism among nuclei consistent with homokaryotic hypothesis. Likewise, homokaryosis has been again challenged by Beaudet et al. (2015) who demonstrated that Next Generation Sequencing underestimates intra-isolate polymorphism that is a strong bias. Overall, some recent publications support the hetrokaryosis hypothesis (Boon et al. 2015; Beaudet et al. 2015).

1.3 AMF biodiversity and interaction with plants

Primarily, observation of about 225,000 terrestrial plant species colonized by approximately 160 species of AMF (Smith 1997) suggested AMF low host specificity. In contrary, some researchers have claimed a high rate of host specificity and local diversity of certain taxa (Yang et al. 2012). If AMF possesses low host specificity, the researchers speculate, the fungi would distribute widely although most AMF have a restricted geographical distribution, which suggests host specificity (Yang et al. 2012). This speculation would indicate that the effects of ecosystems (biogeographical territories and climatic zones) on AMF distribution might be mediated by plants through host selectivity and preference. It has been reported by Alkan *et al.* (2006) that the interface among several AMF and host plant roots would offer multiple benefits to the host plant

in comparison to a host plant colonized by a single AMF species. Concisely, different AMF species might affect plant growth and biodiversity in different ways (Ravnskov and Jakobsen 1995) and possess a defined role in the community structure (Alkan et al. 2006). Thus, appropriate management of mycorrhizae in agriculture should result in a considerable drop in harmful chemical usage and also production expenses. The first step toward AMF application in sustainable agriculture and industrial treatment is a development of precise identification and quantification procedures.

1.4 AMF identification

Arbuscular mycorrhizal fungi (AMF) have traditionally been identified by various taxonomic characters including spore size, colour, sporocarp structure (spore ontogeny, Franke and Morton 1994) and hyphal attachment morphology. Furthermore, the intraradical structures allow identification to the family level the absence of spores (Clapp et al. 1995). In other words in the absence of sexual reproduction, AMF were characterised mainly by morphological criteria. However, there are several limitations for morphological identification and characterisation. For example, AMF sporulation is influenced by ecological conditions and physiological parameters. Several characteristics such as immaturity, degradation, or infection by parasites and chemical reactions might influence spore shapes (Redecker 2002). Meanwhile, similarity of AMF intraradical structures, obscuring usage of the light microscopy method for simultaneously detecting of more than one AMF species in a common root fragment (Alkan et al. 2006). Thus, neither morphology nor ultrastructure characters is sufficient to statistically support phylogenetic inferences for identification and quantification of AMF. Therefore, molecular techniques are

powerful approaches that complement and overcome limitations of morphological-based identification, detection and quantification of AMF providing information on the biodiversity, community structure, and function of AMF in an ecosystem. Many investigations have attempted to develop molecular markers based on nuclear genes in AMF. The high sequence variation within the nuclear genome of AMF (heterogeneity) with high intra-isolate genetic diversity (Kuhn et al. 2001) encountered researchers to some errors and complexities in identification. Existence of multiple sequences within a single AMF isolates or even single spores (Clapp et al. 1995, Sanders et al. 1995), causes overestimation of the number of species making an ambiguity in population analyses interpretation. An ideal molecular marker should be present in a broad range of organisms from the target taxon and easily amplifiable by PCR. The marker should also comprise highly conserved as well as variable regions so that it can be used for phylum to species or even to isolate identification. In AMF, nuclear ribosomal DNA such as the internal transcribed spacer (ITS) of rDNA region including the 5.8S rRNA gene (White et al. 1990, Sanders et al. 1995, Wubet et al. 2003, Hempel et al. 2007), small subunit (SSU) rRNA gene (Wubet et al. 2003, Helgason et al. 1999, Lee et al. 2008) and the large subunit (LSU) rRNA gene (Gollotte et al. 2004, Pivato et al. 2007, Rosendahl et al. 2009, Stockinger et al. 2009) are the most commonly regions used for marker development. Some attempts have been made to develop markers in functional genes such as actin and elongation factor 1-α (Helgason et al. 2003, Sokolski et al. 2010), β-tubulin (Msiska and Morton 2009), phosphate transporter genes (Sokolski et al. 2011), H+-ATPases (Requena et al. 2003). Although the molecular data have changed the systematic by adding several new genera and families (Walker et al. 2004, Walker et al. 2007) there is still a long way to make these approaches trustworthy. For example, although the SSU rDNA region provides the largest taxon sampling for AMF, allowing phylogenetic resolution no more than genus level (Walker et al. 2007). Stockinger et al. (2009) used a combination of SSU-ITS-LSU

rDNA amplicon and resolved AMF phylogenetic down to species' level. Single-copy genes are another type of sequences that have been studied so far to elucidate evolutionary history of AMF (Helgason et al. 2003, Stukenbrock and Rosendahl 2005, Msiska and Morton 2009, Sokolski et al. 2010). Single-copy genes have the advantage that any sequence variation within a single spore can be recognized as a variation among nuclei (Helgason et al. 2003). However, a weak point of analyzing a single gene for evolutionary analyses is that a single-gene tree may not reflect the entire organisms' evolutionary history. In fact, the evolutionary history and divergence time of different segments of the genome may differ from each other and thus, their separate histories do not reflect the entire organisms' evolutionary history (species tree) (Aguilera et al. 2008).

1.5 Anastomosis in AMF

Hyphal fusion (anastomosis) is a phenomenon that results in the cytoplasmic and genetic connection between the same or genetically divergent individuals and is ubiquitous in many filamentous fungi (reviewed in (Saupe 2000, Esser 2006). This phenomenon has crucial functions specifically in the absence of the assumed sexual AMF mode of reproduction. Anastomosis has a major effect in extending the sphere of the underground network among different host plants. It is believed that fitness of AMF and consequently host plants can be enhanced via migration of nutrient, cytoplasm, organelles and even nuclei especially under stressed conditions. Although the host-plant communities influence the frequency of AMF hyphal fusion, their genome similarity is another important prerequisite for anastomosis occurrence (Giovannetti et al. 2001). It is still unknown how the process of hyphal fusion is regulated in AMF. Another important issue is to comprehend the genetic distance required prohibiting the formation of an anastomosis

between two individual AMF, which might lead to a definition for AMF species. The self-recognition and non-self recognition during hyphal fusion were first reported by Giovannetti et al. (1999) in six different species of AMF. Hyphal fusion was observed within the same germline (i.e. a single germinated spore) and between different germlines belonging to the same *Glomus* species (Giovannetti et al. 1999). The authors also detected the exchange of nuclei in the middle of hyphal bridges connecting mycelia (anastomosis) from either the same or different plant species. They introduced anastomosis as a mechanism for maintaining the genetic diversity in the absence of sexual recombination (Giovannetti et al. 2001). The authors performed another test for vegetative compatibility to study the hyphal interactions between six isolates of *G. mosseae* sampled from various worldwide ecosystems and they observed incompatibility responses among them (Giovannetti et al. 2003). However, based on a new report from Croll et al. (2009a) hyphal fusion (anastomosis) can occur even between genetically distinct isolates of *Glomus irregulare* but with lower rates in comparison with genetically similar isolates. This event causes the creation of more genetically diverse networks via nuclei exchange. The authors analyzed the progenies by genetic markers and reported the evidence for biparental inheritance. Sets of experiments have been also conducted by Angelard et al. (2010), which declared both genetic exchange and random distribution (segregation) of nuclei in offspring affecting their phenotypes, plant symbiosis and productivity. The most important consequence of non-self fusion is the formation of heterokaryon, which is in agreement with the hypothesis of heterokaryosis vs. homokaryosis. It seems that in contrast to what is prominent about the fitness of organisms in favour of homokaryosis (Ballard and James 2004) (with a good match between nucleus and mitochondria), mito-nuclear interaction in AMF is under different mechanisms that needs to be explored.

1.6 Mitochondrial DNA and its evolution

Mitochondrial DNA (mtDNA) as it stands from the term, is the DNA located in organelle called mitochondria. mtDNA contains set of genes essential for normal mitochondrial function that is conversion of food chemical energy into a form that cells can consume (adenosine triphosphate, ATP) via oxidative phosphorylation pathway. The current most popular theory about the origin and evolution of mitochondria (endosymbiont theory) suggests that the origin of nuclear genome of eukaryotic cells was occurred in parallel to the origin of mitochondrial genome (Gray et al. 1999, Lang et al. 1999). This is one of the reasons why mtDNA has been studied extensively in many diverse eukaryotic lineages and evolutionary analysis. Mitochondrial genomes are present in multiple copies within a cell that results in better yield of amplification-based methods (e.g., PCR) compared to single nuclear genes. mtDNA was thought to be fully inherited from maternal lines (clonal), and so there is usually no or minor change in mtDNA from parent to offspring with low chance of recombination (Birky 2001). Uniparental inheritance of mitochondria is the main support of the homoplasmy (i.e. all the mitochondrial genomes in the cell are essentially identical), which drastically decreases the chance of recombination (Lynch 1996). However, clonal and uniparental inheritance have been questioned and rejected in some organisms such as mussel from bivalve family (Breton et al. 2007). Beside this phenomenon, mtDNA higher rate of mutation to compare with nuclear DNA (because of biochemical processes of mitochondrion itself, lack of histones in DNA structure, and low efficiency of the DNA repair system (Rand 2001)), suggested this genomic pool as a potential candidate tool for evolutionary analysis. mtDNA elevated mutation rate made it highly variable in natural population offering a good tracer for population history over a short period of time. It has also been thought that the

mutation rate in mtDNA was constant, and mtDNA is almost neutral which both have been questioned by collection of recent data (Galtier et al. 2009). Therefore, caution should be taken, and mtDNA should be studied in a variety of organisms to upgrade the understanding of its evolutionary constraints.

1.7 Mitochondrial inheritance in AMF

Mechanisms of mitochondrial inheritance in AMF remains to be demonstrated. So far, the putative homoplasmy reported for the mitochondrial genome of Rhizophagus irregularis could be explained by four hypotheses: (1) Vegetative incompatibility among genetically different individuals prevents co-inheritance of divergent mtDNA haplotypes in the progeny. (2) The exchange of both mitochondrial haplotypes occurs, but homoplasmy is reestablished in subsequent generations. (3) Homoplasmy existed only in AMF *in vitro* cultures because of a homogenized condition, and it does not exist in the nature. (4) Based on a hint by Barr et al. (2005), heteroplasmy might rarely or only transiently occur, leading to recombination and removal of some mutations. After all, in case of production of mitochondrion harboring incomplete genome, mitochondrial autophagy (mitophagy), results in a selective degradation of damaged mitochondria in cells. Therefore, studying the inheritance pattern, maintenance, and stability of mtDNA in AMF help to address these important biological questions and facilitate the development of mitochondrial markers as a promising choice for identification purposes.

1.7 Mitochondrial inheritance in AMF

In order to circumvent depletions and the probable complexity of nuclear genes for AMF identification, studies of an independent genomic pool of mtDNA have noted several advantages (Moritz et al. 1988, Raab et al. 2005, Borstler et al. 2008, Lang and Hijri 2009, Lee and Young 2009). Conversely, the use of such genomic information is not without pitfalls because we have little understanding of selective forces that cause mitochondrial gene loss and reshuffling. Bruns et al. (1989) found that the gene for the large subunit of the mitochondrial ribosomal RNA (rnl gene) is useful for molecular identification of the ectomycorrhizal fungi. It has been reported that the rnl gene has a very low degree of intra-isolate genetic variation in the AMF *G. intraradices* (*G. irregulare*) and *G. proliferum* (Raab et al. 2005). However, sequences of this region were found to be polymorphic among isolates of these species (Raab et al. 2005), and sequence variation is more substantial (Borstler et al. 2008). It has also been reported that both exons and introns of the rnl offered opportunities to develop markers for discriminating haplotypes of *G. intraradices* (*G. irregulare*). Subsequently, investigating and mapping of several mitochondrial genomes from R. irregularis species proposed homoplasmy in AMF offering mitochondrial genome as a promising choice for species identification (Lee and Young 2009, Formey et al. 2012). Rapid genetic segregation (i.e. random distribution of mtDNA in the offsprings) (Lee and Young 2009) and/or an active process of transmitting homogenous mtDNA to descendants (Ling and Shibata 2004) are suggested to be the main reasons for lack of polymorphism in AMF mtDNA. Thus, the mitochondrial genome of AMF has a yet-to-be-explored potential for understanding the evolution of this important fungal group and designing molecular tools for further characterization in natural populations.

1.9 Research objectives, hypotheses and thesis presentation

When I started my Ph.D. project on AMF, little was known about their mtDNA structure, characteristics, evolutionary constraints and signals. There were only one published mitogenome (*Glomus intraradices* isolate#494, Lee and Young, 2009) paper and few studies based on ribosomal large subunit (mtLSU) (Raab et al. 2005; Borstler et al. 2008)! Considering the difficulties and complexities of AMF identification required for their application, many questions were raised regarding the sue of mtDNA for the development of molecular toolkits for identification, quantification in order to investigate evolution, function, population genetics and ecology of AMF. In this context, the general objective of my Ph.D. project was to investigate mitochondrial genome diversity and their evolutionary mechanisms among representative glomeromycotan taxa using next-generation sequencing, bioinformatics and molecular biology approaches.

The specific objectives are:

1. To investigate mitochondrial genome evolution using comparative mitogenomic analyses of closely related species and isolates as well as phylogenetically distant taxa of AMF
2. To assess mitogenomes and mt genes for marker development and phylogenetic analyses.
3. To explore mitogenomes, inheritance among compatible isolates of the model AMF *Rhizophagus irregularis* through anastomosis formation

I have addressed the following scientific questions:

i. What is the mt genome structure, gene synteny in *Gigaspora rosea*, a phylogentetically distant AMF taxon compared to *R. irregularis* DAOM 197198?

ii. What is the polymorphism status of mtDNAs among isolates of the same species?

iii. Is mt gene synteny conserved among taxa of the genus *Rhizophagus*?

iv. Can mtDNA be used for development of molecular toolkits and for phylogenetic analysis in AMF?

v. Do the mt genes possess the same evolutionary signal?

vi. Are mtDNA haplotypes from two crossed isolates of *R. irregularis* inherited to the progeny after anastomosis?

vii. Does heteroplasmy as an outcome of anastomosis make drawbacks in utilization of mtDNA for phylogenetic and identification analyses?

I tested the following hypotheses in my Ph.D. project:

a. mtDNA evolves rapidly in AMF.

b. mtDNA is highly polymorphic among isolates of a given species.

c. mtDNA is suitable for developing isolate specific markers.

d. mt genes are not possessing the same evolutionary signal as it was thought and so some mt genes should be targeted for facilitating phylogenetic analyses.

e. mtDNAs are exchanged via anastomosis but there are some unknown mechanisms (e.g., segregation and selection) which act in favor of homoplasmy.

Through the first investigation (2nd chapter), mtDNA of *Gigaspora rosea* DAOM 194757 has been sequenced revealing relatively large genome size (97,349 bp) comparing other sequenced *Glomus irregulare* who's name has been changed to *Rhizophagus irregularis*, during my Ph.D. project. We also annotated unorthodox fragmented genes (*rns* and *cox1*) in the genome of *Gigaspora rosea*.

The 3rd chapter of my thesis (second article) consisted of sequencing, assembling and analyzing mtDNA of *Rhizophagus* sp. DAOM 229456. Comparative mitogenomics on the

mitochondrial genome of *Rhizophagus* sp. DAOM 229456 revealed the first evidence of AMF interspecific exchange of mitochondrial coding sequences resulting in formation of hybrid genes in *atp6*, *atp9* (coding for the subunit 6 and 9 of the ATP synthase complex), *cox2* (cytochrome C oxidase subunit 2) and *nad3* (NADH dehydrogenase subunit 3) genes.

MtDNA sequencing and mapping of *Rhizophagus* sp. DAOM 213198 also reveal an a novel mtDNA structure that has never been reported in fungi. This peculiar organization has been characterized in the 4th chapter (third article). Further quantification approaches have been conducted in order to measure mitochondrial genomes copy number. Comparative mitogenomics analyses were also applied to trace the probable origins and causes of this novel mtDNA structure formation in *Rhizophagus* sp. DAOM 213198.

Finally in 5th chapter (fourth article), I evaluated the power of individual mt genes in phylogenetic analyses of AMF and revealing their evolutionary relationship. Mitochondrial genes evolutionary signals have been compared to phylogenetic signal of "supergene" set signal in order to explore the best-performing genes. Mitogenomics comparative analysis also has been implemented to identify uncertain position isolate of AMF and to reveal the evolutionary history among closely related isolates.

The 6th chapter (fifth article) of my thesis represents the investigation of heteroplasmy status within genetically divergent isolates of *R. irregularis* in case of anastomosis occurrence. We performed three crossing combinations both in pre-symbiotic and symbiotic phases. Progeny spores per each crossing combination were genotyped using isolate-specific markers that have been developed in the mitochondrial genome. Genotyping patterns of individual spores from the progenies revealed the fate of the two parental mtDNA haplotypes. Further germination of some progenies and the genotyping pattern of them also gave us a speculation about the fate of mtDNA in next generation and persistence of our result.

My thesis is presented under the scientific article form for the Doctorate in Biological Science Program at the Université de Montreal. Chapter 1 of the thesis (Introduction and Literature review) introduces the current knowledge about the arbuscular mycorrhizal fungal biology, genetics and genomics as well as mitochondrial genome evolution. The personal experiments, methods and results are then introduced in the chapters 2, 3, 4, 5 and 6 in four published articles and one submitted manuscript. Chapter 7, the last section of the thesis, serves as a general discussion and conclusion of all the obtained results following by the perspectives of this project.

Presentation of article 1

In 1990, arbuscular mycorrhizal fungi were organized in order "Glomales" containing three families (*Acaulosporaceae, Gigasporaceae,* and *Glomeraceae*) and six genera (*Acaulospora*, *Entrophospora*, *Gigaspora, Glomus*, *Sclerocystis,* and *Scutellospora*) within phylum Zygomycota (Morton 1990). However, the authors have not implemented molecular aspects and evidences such as their symbiotic habit, lack of zygospores that suggest AMF form a monophyletic group distinct from other Zygomycotan lineages (Schüβler et al. 2001). Based on these evidences, the phylum Glomeromycota has been proposed by Schüßler et al. (2001). Phylogenetic analyses based on 18S rDNA revealed that Glomeromycota are the sister group to Ascomycota and Basidiomycota (Schüßler et al. 2001). Advances made in fungal genome sequencing last decade, allowed generating a huge datasets in fungal mtDNAs such as many mtDNA of the Dikarya (Basidiomycota and Ascomycota) as well as basal fungal lineages Chitridiomycota, Zygomycota and Glomeromycota. Publication of the first mitochondrial genome of AMF species by Lee and young in 2009 opens the avenue through mitochondrial assessment of AMF (Lee and Young 2009). The increasing number of sequenced mtDNAs in fungal kingdom offers opportunities to infer robust phylogenies using phylogenomics approaches. Sequencing of an AMF member relatively distinct from the available sequenced mtDNAs (*Rhizophagus* and *Glomus* members) would enhance our knowledge of AMF evolutionary history.

In this article, we report the complete mtDNA sequence of *Gigaspora rosea*, which encodes two fragmented genes, transcripts of one of which undergo group I intron–mediated trans-splicing. We further report the results of a phylogenetic analysis of mitochondrial proteins that allows more confident positioning of Glomeromycota within Fungi.

This article was published in Molecular Biology and Evolution. I have contributed significantly to data analyses using bioinformatics in particular analyses of detailed fragmentation annotation (via generating secondary structure) for *G. rosea*. I have also contributed together with D. Beaudet and L. Forget to RNA analyses using molecular biology methods. I have written some sections of this article.

Chapter 2 - Group I intron–mediated trans-splicing in mitochondria of *Gigaspora rosea* and a robust phylogenetic affiliation of arbuscular mycorrhizal fungi with Mortierellales

Maryam Nadimi[*1], Denis Beaudet[*1], Lise Forget[2], Mohamed Hijri[1], and B. Franz Lang[+2]

[1] Université de Montréal, Département de sciences biologiques, Institut de recherche en biologie végétale (IRBV), 4101 rue Sherbrooke Est, Montréal, QC, H1X 2B2, Canada.

[2] Département de Biochimie, Centre Robert-Cedergren, Université de Montréal, Québec, Canada

*These authors contributed equally to this work.

Published in : Molecular Biology and Evolution (MBE), March 12, 2012. 29(9): 2199-210, doi: 10.1093/molbev/mss088.

2.1 Abstract

Gigaspora rosea is a member of the arbuscular mycorrhizal fungi (AMF; Glomeromycota) and a distant relative of *Glomus* species that are beneficial to plant growth. To allow for a better understanding of Glomeromycota, we have sequenced the mitochondrial DNA of *G. rosea*. A comparison with *Glomus* mitochondrial genomes reveals that Glomeromycota undergo insertion and loss of mitochondrial plasmid-related sequences and exhibit considerable variation in introns. The gene order between the two species is almost completely reshuffled. Furthermore, *Gigaspora* has fragmented *cox1* and *rns* genes, and an unorthodox initiator tRNA that is tailored to decoding frequent UUG initiation codons. For the fragmented *cox1* gene, we provide evidence that its RNA is joined via group I-mediated trans-splicing, whereas *rns* RNA remains in pieces. According to our model, the two *cox1* precursor RNA pieces are brought together by flanking *cox1* exon sequences that form a group I intron structure, potentially in conjunction with the *nad5* intron 3 sequence. Finally, we present analyses that address the controversial phylogenetic association of Glomeromycota within fungi. According to our results, Glomeromycota are not a separate group of paraphyletic zygomycetes but branch together with Mortierellales, potentially also Harpellales.

2.2 Key words

Arbuscular mycorrhizal fungi (AMF), mitochondrial genome, intron evolution, phylogeny, tRNA structure, genetic code.

2.3 Introduction

Arbuscular mycorrhizal fungi (AMF) is a group of ubiquitous soil-borne fungi that form symbiotic associations with the majority of vascular plants (Parniske 2008). AMF are obligate

biotrophs, that is, they are unable to grow without a host plant that provides them with carbohydrates; in turn, AMF transfer nutrients such as phosphate to the plant (reviewed in (Strack et al. 2003)). At the cellular level, AMF are characterized by the formation of large, multinucleate hyphae, and asexual spores (e.g., (Marleau et al. 2011)). Apparently, the genetic segregation of the hundreds of distinct nuclei that are present in these species does not follow canonical but rather population rules, and recent analyses demonstrate substantial sequence variation in certain nuclear genes (Hijri and Sanders 2005, Croll et al. 2009, Croll and Sanders 2009, Boon et al. 2010). It is therefore no surprise that the *Glomus irregulare* nuclear genome project has turned into a sequence assembly nightmare (Martin et al. 2008). Yet in stark contrast, the first complete *Glomus* mitochondrial DNAs (mtDNAs) that have been deciphered recently by 454 sequencing are homogeneous in sequence (Lee and Young 2009) and GenBank ′FJ648425; Bullerwell CE, Forget L, Lang BF, unpublished data), that is, genetic segregation of mtDNA is as effective in Glomus as in other fungi. In these two cases, long homopolymer stretches that introduce systematic pyrosequencing error are surprisingly absent. In other, more A+T-rich mtDNAs, however, we have observed intolerable levels of 454 sequence error (close to one per 1 kbp sequence on average in a heterolobosean amoeba; Bullerwell CE, Forget L, Lang BF, unpublished data) causing frameshifts in several protein-coding genes. It therefore remains advisable to carefully examine homopolymer-rich sequences for potential error, for instance by resequencing with Sanger technology.

The taxonomic and phylogenetic identity of AMF have been, and continues to be, controversial. Initially assigned to zygomycetes, a fungal taxon that is strongly suspected to be paraphyletic (e.g., (Schwarzott et al. 2001, Seif et al. 2006, Hibbett et al. 2007, Liu et al. 2009), AMF have been recently moved into a separate fungal phylum, Glomeromycota (Hibbett et al. 2007). Yet the underlying published phylogenies are controversial and often lack significant

statistical support, either due to a limited amount of sequence data (based on one or only few gene sequences), poor taxon sampling, or a combination of both. In many instances, analyses further suffer from potential phylogenetic artifacts such as long-branch attraction (e.g., Felsenstein 1978).

A phylogenetic data set of complete mtDNA sequences is currently restricted to two *G. irregulare* isolates (Lee and Young 2009 and GenBank ′FJ648425) having identical coding sequences. Likewise, a previous phylogenomic analysis with a large number of nuclear genes (Liu et al. 2009) had only limited taxon sampling. Accordingly, phylogenetic analyses with both mitochondrial and nuclear genome data have provided only a tentative answer to the question of where AMF belong within Fungi. In some cases, they both show a weak affinity of Mortierellales with Glomeromycota (Lee and Young 2009, Liu et al. 2009), unsupported however by strict statistical analysis (such as the AU test, Shimodaira 2002). An updated, comprehensive fungal phylogenomic analysis with nuclear sequence data published in 2011 (Ebersberger et al. 2012) nicely summarizes the confusing state of the art, commenting that ‘‘at the moment, available data do not allow to confidently attach glomeromycetes to the phylogenetic backbone of the fungi.’’ In the latter phylogenomic analysis, Mortierellales are shown separate from Mucorales, that is, excluding Mortierellales from a monophyletic taxon Mucoromycotina favored by others and in contradiction to conclusions reached in a previous phylogenomic analysis (Liu et al. 2009). Evidently, better taxon sampling of genomic data sets is required to resolve these questions, in particular by adding to both mitochondrial and nuclear gene data sets members of Mortierellales, and AMF lineages that are distant from Glomeraceae.

Our rationale for sequencing the *Gigaspora rosea* mtDNA is that Gigasporaceae are at a large evolutionary distance to *Glomus* species, with clearly distinct morphological characteristics. Gigasporaceae form auxiliary cells in the extraradical mycelium, and giant spores that are usually

larger than 200 lm and visible to the naked eye. These atypical spores are formed individually from hyphae and contain funnel-shaped hyphal attachments that extend from a specialized bulbous sporogenous cell. Another motive for sequencing mtDNAs from additional Glomeromycota relates to the origin and distribution of introns, other endonuclease-based mobile elements, and mitochondrial plasmid-like inserts in mtDNAs that are frequent in *Glomus* species ((Lee and Young 2009); Lang BF, unpublished data). In fact, one of the group I introns of a glomeromycotan cox1 gene might have been transmitted to plant mitochondria (Vaughn et al. 1995, Adams et al. 1998, Seif et al. 2005, Lang and Hijri 2009), potentially as part of symbiotic plant–AMF interactions. Group I introns are mobile genetic elements that may insert into intron-less gene copies (i.e., with respect to a given intron insertion point) by a process called intron homing (Colleaux et al. 1986, Lambowitz and Zimmerly 2004). They are further characterized by an RNA that folds into a distinct secondary structure, consisting of up to nine base-paired helical regions (P1–P9) that are involved in bringing respective exons into close proximity and in the splicing reaction itself (Anziano et al. 1982, Michel et al. 1982, Waring et al. 1982). In many but not all cases (Lang et al. 2007) group I introns are self-splicing in vitro, that is, the RNA is a ribozyme capable of catalyzing its own excision from precursor RNA (Cech et al. 1983, Jacquier and Rosbash 1986, Schmelzer and Schweyen 1986, van der Veen et al. 1986).

Intron splicing usually occurs in cis, but a few cases of trans-splicing are known, and we provide evidence in this paper that this also applies to a *G. rosea* mitochondrial gene. For organelle genes, trans-splicing is usually mediated by group II introns (for a review, see (Bonen 2008)) but in a few recently discovered instances also by group I introns (Burger et al. 2009, Grewe et al. 2009, Pombert and Keeling 2010, Hecht et al. 2011). Curiously, all known trans-splicing group I introns are located in *cox1*, with identical insertion points in the *quilwort Isoetes* and the *spikemoss Selaginella* but also in the evolutionarily distant *Gigaspora* and the metazoan

Trichoplax (supplementary fig. S5.2.1, Supplementary Material online). Trans-splicing entails the joining and ligation of discontinuous coding regions that are transcribed separately, that is, located on distinct RNA molecules. For group I and II introns, these separate transcripts may be brought together by interaction of partial (i.e., fragmented) intron RNA sequences that fold into the typical RNA structure that allows splicing to take place. In more complicated, rare instances, the flanking sequences of precursor transcripts alone do not code for the full intron structure, but additional helper RNAs are involved. In the two known instances, three distinct RNA molecules in trans are required to form a complete intron RNA structure (Goldschmidt-Clermont et al. 1991, Knoop et al. 1997). In any case, claims for intron-mediated trans-splicing should be based on evidence for 1) the presence of a mature RNA that is ligated exactly at predicted exon–intron junctions (e.g., based on multiple sequence alignments) and 2) an inferred intron RNA structure that is complete, matching previously established conservation rules perfectly. Note that not all discontinuous genes require trans-splicing. For instance, known fragmented rRNAs remain in separate transcripts, folding into a functional ribosomal structure, but without the requirement for trans-splicing (e.g., (Boer and Gray 1988, Schnare and Gray 1990).

In this article, we report the complete mtDNA sequence of *G. rosea*, which encodes two fragmented genes, transcripts of one of which undergo group I intron–mediated trans-splicing. We further report the results of a phylogenetic analysis of mitochondrial proteins that allows more confident positioning of Glomeromycota within Fungi.

2.4 Materials and Methods

2.4.1 Fungal Material

Spores of *G. rosea* (DAOM194757) were extracted from soil of *in vivo* pot cultures with leek

as a host plant, using a wet-sieve cascade (400, 250, and 60 μM). Spores were further purified by centrifugation through a water/50% sucrose step gradient (50-ml tube; 5 min at 5,000 rpm in an Eppendorf 5804 centrifuge; (Esch et al. 1994). The resulting spore containing layer from the upper interface of the 50% sucrose solution was collected, and clean spores were sorted manually under a binocular.

2.4.2 DNA Purification

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

2.4.3 RNA Purification

Freshly harvested *G. rosea* spores from *in vivo* pot cultures were vortexed with 1 ml TRIzol reagent (Invitrogen) and glass beads for 10 min. Other RNA extraction steps followed the manufacturer's protocol. Total RNA was then purified with the RNeasy mini kit (Qiagen), treated with DNase using the Turbo DNA free kit (Applied Biosystem), and stored at -20 °C. RNA quality was checked by electrophoresis on a 1.5% agarose gel, stained with ethidium bromide, and visualized under UV light.

2.4.4 Reverse Transcriptase–Polymerase Chain Reaction

We have used reverse transcriptase–polymerase chain reaction (RT-PCR) to test for intron excision, trans-splicing, and the presence of (apparently nonintron) sequence inserts in rRNAs. In all experiments, PCR amplifications were performed either with or without prior conversion of RNA into cDNA (i.e., addition of RT). As an additional control, genomic DNA was PCR amplified. To test for potential trans-splicing of *cox1* and *rns*, primers were designed within the coding regions that flank predicted break points. In case of *cox1*, primers were located in *cox1* exon 4 (5′-CTGTATTGGTCACTGCCGT-3′) and exon 5 (5′-AAAGGCTGAAATAACATGGCT-3′). For *rns*, the two respective primers were 5′-ACCTTGATCCAGCCAACTAGA-3′ and 5′-CACACTATCGTATCTCAGCGTC-3′. To test for functionality of the RT-PCR assay (positive control), an internal region of *cox1* exon 4 was amplified (primers 5′-CTCTAGCAGGGACCCAGTC3′ and 5′ -CCGGATCATAGAAGCAGGT3′). To verify whether several inserts in *rnl* and *rns* remain at the RNA level (suspected because of an apparent lack of conserved intron RNA secondary structure), primers were designed within the flanking regions of these inserts. For *rns*, the primers were 5′-GTGCATTGTCATCACAGGTG-3′, 5′-CGAGTTACAGAGCACAGTTCG-3′, and 5′-GTCTCGTAACAAGCCTCCTTAAC-3′ (the latter is located inside the predicted insert). In the case of *rnl*, five primers were designed to test inserts 3, 4, and 5 (5′-AAGTAGAGGCTCCAGAAGCAG-3′, 5′-GAAGGGTTTCATGAGTAAGGTGA-3′, 5′-CAGGTCTGCAGGTTCACG-3′, 5′-CGTTCAGTCTTAACACTTGGC-3′, and 5′-CCCCTTTTTAGTGCCGC-3′). We further verified that the two predicted *rnl* introns are indeed excised at the RNA level, using primer (5′-GGGCGCGTCTGTTTACTTA-3′) upstream of the first intron (group II), in conjunction with a primer downstream of the second group IB intron (5′CGAGTACCGGTACCAGAGTAGGT-3′). For samples undergoing cDNA synthesis, relevant

primers (0.1 µl of 5 µM stock solutions), 2 µl RNA solution (~1 to 10 ng), and 8 µl of RNase-free water were mixed, denatured at 75 °C for 2 min, and immediately placed on ice for 2 min. Then, 4 µl RT buffer, 2 µl of each dNTP (10 mM stocks), 4 µl water, and 0.1 µl (1.5 U) AMV RT enzyme (Roche Applied Science) were added. After incubation at 45 °C for 45 min in the reaction mix, the RNA was denatured again at 75 °C for 2 min, placed on ice, fresh AMV RT enzyme was added, and the samples incubated for another 45 min at 45 °C. Reaction products were stored at -20 °C until use. Subsequent PCR amplifications were carried out with the Expand High Fidelity PCR system (Roche Applied Science) in a total volume of 40 µl, containing either cDNAs (4 µl) or positive and negative controls (0.4 µl RNA or 0.2 µl genomic DNA), primers (4 µl each from 5 µM stocks), 4 µl PCR buffer (10X, without MgCl2), 3.2 µl MgCl2 (25 mM), 0.4 µl (1.5 U) DNA polymerase, 4 µl dNTPs (2 mM), and 1 µl RNase (5 µg/ml). Cycling parameters were 4 min at 95°C, followed by 33 cycles of 15s at 95°C, 20 s at a 53 to 57 °C temperature gradient, 20 s at 72 °C, and a final elongation step of 4 min at 72 °C. PCR products were separated by electrophoresis in a 1.5% (w/v) agarose gel and visualized with ethidium bromide under UV light.

2.4.5 Sequencing, Assembly, and Gene Annotation

Gigaspora rosea total DNA was sequenced by 454 shotgun technology, and the resulting 587,881 reads were assembled with Newbler (Genome Quebec Innovation Center, McGill University, Montreal; Titanium Flex, 1⁄2 plate). The resulting 26 mitochondrial contigs were combined into a single circular-mapping DNA by PCR amplification of total DNA and Sanger sequencing of the PCR fragments. Regions with potential 454 sequence error (due to homopolymer motifs) were resequenced in the same way. Gene annotation was performed with

MFannot (http://megasun.bch.umontreal.ca/cgi-bin/ mfannot/mfannotInterface.pl), followed by manual inspection and addition of missing gene features. Mfannot predicts group I and II introns, tRNAs, RNase P RNA, and 5S rRNA with Erpin (Gautheret and Lambert 2001) as a search engine, based on RNA structural profiles developed in house. Exons of protein-coding genes are inferred in a first round with Exonerate (Slater and Birney 2005) and then for less well-conserved genes with HMM profile searches (Eddy 2008; based on models for all known mtDNA-encoded proteins). Miniexons that are not recognized by Exonerate and that may be as short as 3 nucleotides are inferred by the presence of missing conserved protein regions (with reference to multiple protein alignments including other species) and orphan introns. The precise placement of small exons is based on the best fit of HMM protein profiles and on the fit with conserved nucleotide sequence profiles of respective (group I or II) exon–intron boundaries. Genes encoding the small and large rRNA subunits are predicted with HMM profiles that cover the most highly conserved domains, allowing precise predictions of the small subunit rRNA termini but only approximate positioning of large subunit rRNA ends. The latter termini, as well as the precise exon–intron structure of rRNA genes, are predicted manually (using comparative structure modeling principles). Furthermore, RT-PCR experiments are evaluated to distinguish between true introns and sequence insertions that are not removed by splicing. In any case, automated annotations are curated manually to account for MFannot warnings (e.g., potential *trans*-spliced genes, gene fusions, frameshifts, alternative translation initiation sites, failure to identify miniintrons, etc.), to correct potential errors, and to find features that are not recognized by the automated procedures. The annotated sequences (mtDNA and plasmid) are available at GenBank (JQ693395 and JQ693396).

2.4.6 Phylogenetic Analysis

The data set contains 13 mitochondrion-encoded proteins (*cox1*, *2*, *3*, *cob*, *atp6*, *9* and *nad1*, *2*, *3*, *4*, *4L*, *5*, *6*) and includes sequences from all zygomycetes for which complete mtDNA sequences are available. Protein collections were managed and automatically aligned, trimmed, and concatenated with Mams (developed in house; Lang BF, Rioux P, unpublished data). Mams uses Muscle (Edgar 2004) for an initial alignment, followed by a refinement step with HMMalign (S. Eddy; http:// hmmer.janelia.org). The final data set contains 37 taxa and 3,664 amino acid positions and is available from the authors upon request.

Phylogenetic analyses were performed by Bayesian inference (PhyloBayes; (Lartillot and Philippe 2004)) using the CAT+Gamma model, six discrete categories, four in- dependent chains, 10,000 cycles (corresponding to ~550,000 generations), and the -dc parameter to remove constant sites. The site removal reduces the total number of amino acid positions to 3048. The first 1,700 cycles were removed as burn-in. The robustness of internal branches was evaluated based on 100 jackknife (60%) replicates, as modeling of duplicated sequence sites generated by bootstrap analysis is problematic with the Bayesian approach. Maximum likelihood (ML) analysis was performed with RAxML-HPC v7.2.2 (Stamatakis 2006), using the LG model (PROTGAMMALGF), and the fast bootstrapping option (100 replicates).

2.5 Results and Discussion

2.5.1 Comparison of *G. irregulare* and *G. rosea* Mitochondrial Genomes

The mitochondrial genome of *G. rosea* has been completely sequenced by a combination of 454 sequencing from total DNA and filling of the remaining 26 gaps (due to low sequence coverage and 454 sequence error) by PCR amplification and Sanger sequencing. The size of *G.*

rosea mtDNA is relatively large (97,349 bp) when compared with those of the two *G. irregulare* isolates (70,606 and 70,799 bp) and contained several homopolymer stretches that required correction of the 454 data by traditional Sanger sequencing. In addition, paired-end sequencing was performed to avoid misassembly of the genome due to the presence of long sequence repeats. Whenever sequence repeats occur in a genome, more than one genome conformation may exist due to recombination events, but we have no evidence that this occurs in *G. rosea*. The large genome size of *G. rosea* is essentially due to extended intergenic regions, not a variation in the number or size of introns (table 2.1), or the presence of mitochondrial plasmid insertions that are otherwise common in fungi.

Table 2.1. Gene and Intron Content in Selected Fungal mtDNAs.

Species	Genes and Introns										
	rnl, *rns*	*atp6*, *8*, *9*	*cob*	*cox1*, *2*, *3*	*nad1–6*[a]	*trn A-W*	*rnpB*	*rps3*	ORFs[b]	Intron I[c]	Intron II[c]
Gigaspora rosea	2	3	1	3	7	25	0	0	4	13	1
Glomus irregulare 494	2	3	1	3	7	25	0	0	8	24	0
G. irregulare 197198	2	3	1	3	7	25	0	0	8	26	0
Smittium culisetae	2	3	1	3	7	26	1	1	3	14	0
Mortierella verticillata	2	3	1	3	7	28	1	1	7	4	0
Rhizopus oryzae	2	3	1	3	7	23	1	0	4	9	0
Allomyces macrogynus	2	3	1	3	7	25	0	1	4	26	2
Saccharomyces cerevisiae[d]	2	3	1	3	0	25	1	1	3	9	4

[a] Includes *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6*.

[b] Only ORFs greater than 100 amino acids in length are listed, not including intronic ORFs and *dpo* and *rpo* fragments.

[c] Intron I and Intron II denote introns of group I and group II, respectively.

[d] *S. cerevisiae* strain FY 1679 (Foury, Roganti et al. 1998).

The high number of initial sequencing gaps was surprising as according to our experience, the same amount of total 454 sequence information resulted in complete or almost complete coverage of *Glomus* spp. mtDNA. Among possible explanations, a substantial contamination with DNA from foreign species seems unlikely, as spores were carefully cleaned and sorted individually prior to extraction of total DNA. It remains possible either that the copy number of mtDNA is lower in *G. rosea* than in other AMF species, that its nuclear genome is much more complex, or that the spores harbor endosymbionts accounting for a substantial increase in total DNA relative to mtDNA content. Given the available sequence data, we have no reason to favor any of these scenarios. As in most other fungi, the *G. rosea* mtDNA maps as a circular molecule

(fig. 2.1) but is likely organized as a linear multimeric concatamer *in vivo* (Bendich). To allow for easier genome comparison, we have opened the circle at a previously introduced standard position, upstream of *rnl*. Conservation of mitochondrial gene order between *Gigaspora* and *Glomus* is virtually nonexistent and, in contrast to *G. irregulare*, genes are encoded on both strands of the *G. rosea* mtDNA. These genes belong to the basic fungal set (table 2.1), including all required tRNAs. Three tRNAs have a CAU anticodon, one each for initiator and methionine elongator tRNAs and a third one in which a potential lysidine modification of the anticodon C residue would allow decoding of AUA as isoleucine (e.g., (Muramatsu et al. 1988; Weber et al. 1990). Notably, the predicted initiator methionine tRNA exhibits structural features in the anticodon loop that never occur in orthodox tRNAs, entailing a most unorthodox modification of decoding preferences (for more details, see below).

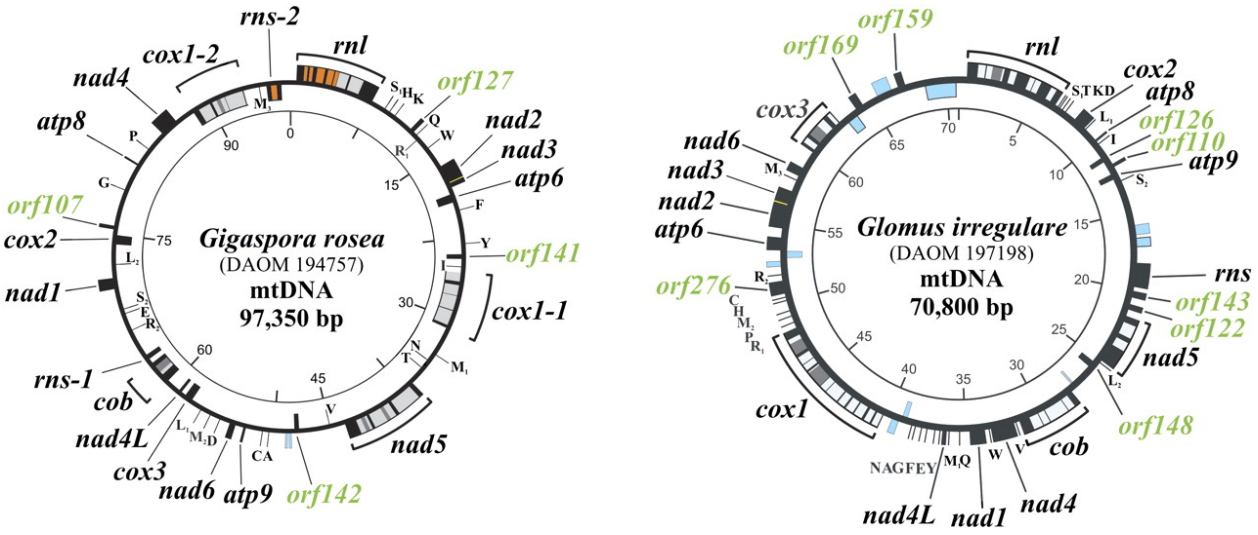


Figure 2.1. Comparison of *Gigaspora rosea* and *Glomus irregulare* mitochondrial genomes.
The circular-mapping genomes of *G. rosea* and *G. irregulare* were opened upstream of *rnl* to allow for easier comparisons. Genes on the outer and inner circumference are transcribed in clockwise and counterclockwise direction, respectively. Arcs indicate coding regions interrupted by introns. Gene fragments in *G. rosea* are numbered −1 and −2. Boxes of coding regions are

filled black, intron ORFs dark gray, and introns light gray. Regions with similarity to mitochondrial plasmid-like DNA polymerase are marked light blue and insertions in *rns* and *rnl* that are not excised from the rRNA (i.e., are not introns) in orange. Gene and corresponding product names are *atp6*, ATP synthase subunit 6; *cob*, apocytochrome b; *cox1–3*, cytochrome c oxidase subunits; *nad1–4*, *4L*, *5–6*, NADH dehydrogenase subunits; *rnl*, *rns*, large and small subunit rRNAs; A-W, tRNAs, the letter corresponding to the amino acid specified by the particular tRNA.

As in *G. irregulare* (Lee and Young 2009), genes for the ribosomal protein *rps3* and the RNA component of mitochondrial RNase P (*rnpB*) are absent, although they occur in most zygomycetes (Seif et al. 2005). However, whereas *rps3* may have been transferred to the nuclear genome, this inference remains more contentious for *rnpB* as it would require targeting of a relative large RNA molecule from the cytoplasm back into mitochondria. An alternative solution would be a nucleus-encoded protein-only RNase P activity that replaces the ribozyme, as it occurs in animal and plant mitochondria (Holzmann et al. 2008; Gobert et al. 2010).

2.5.2 Mitochondrial Plasmid Insertions

Mitochondrial autonomously replicating plasmids are widespread. They may carry their own genes for DNA replication (*dpo*) and transcription (*rpo*), and in some cases, plasmids are known to integrate into mtDNA. There is no biochemical or genetic evidence that complete or fragmented versions of *dpo* and *rpo* are functional when inserted into mtDNA; in fact, plasmid insertion into mtDNAs is followed by rapid genome reorganization and loss (e.g., Bertrand et al. 1986; Vierula and Bertrand 1992; Baidyaroy et al. 2000; Hausner 2011), confirming the view that *dpo* and *rpo* are required for plasmid replication and transcription only.

We have used our shotgun data containing both mitochondrial and nuclear sequence reads to search for plasmid-related sequences. *G. irregulare* mtDNA carries many more and larger plasmid-related DNA polymerase (*dpo*) gene fragments than *G. rosea* (filled light blue boxes; two tiny fragments in *G. rosea* are at map position ;50 kbp; fig. 2.1). In addition, *G. rosea* has an apparently complete, plasmid-like *dpo* (but no *rpo*) in a 3,582-bp contig. Given its relatively high sequence coverage with respect to nuclear genes, this element is likely a freestanding plasmid that only encodes *dpo*. We have not succeeded in demonstrating a circular plasmid version by PCR amplification (including the use of a dedicated long PCR kit; data not shown) and therefore assume that the plasmid is linear. Its cellular location (cytoplasmic or mitochondrial) remains unknown. The *dpo* gene is translated with the standard genetic code, but this would be compatible with both a cytoplasmic and mitochondrial location (for a more detailed discussion of the mitochondrial genetic code, see below). Whatever its cellular location, the *G. rosea* plasmid *dpo* is most closely related to mitochondrial *dpo* fragments, in both *Glomus* and *Gigaspora*, belonging to the family of plasmids that repeatedly invade mitochondrial genomes. A freestanding form of a mitochondrial plasmid has so far not been described for glomeromycotan species.

2.5.3 Evolution of the Mitochondrial Genetic Code in Glomeromycota

Some members of the paraphyletic zygomycete lineages, such as *Rhizopus oryzae*, but also the rapidly evolving *Smittium culisetae*, have retained the standard genetic code for the regular set of mitochondrial protein-coding genes, a trait inherited from their bacterial ancestors (Seif et al. 2005). However, *Mortierella verticillata* reassigns UGA ‘‘stop’’ codons as tryptophan, one each in *nad3* and *nad4*. Likewise, UGA(Trp) codons are also present in the *S. culisetae* group I

intronic open reading frame (ORF) 283 and ORF 248, encoding homing endonucleases of the LAGLI/DADG type (Seif et al. 2005). In Glomeromycota, deviations from the standard genetic code also vary. A few clear-cut instances of UGA(Trp) codons are found in the two published *G. irregulare* mtDNAs but not in *G. rosea*. In turn, translation initiation with either AUG(Met) or GUG(Met) leads to seven proteins in *G. rosea* (i.e., a surprising 50% of standard protein-coding genes) that are severely truncated at their amino-terminal. In these instances, sequence similarity at the protein level extends further upstream to potential UUG initiation codons (supplementary fig. S2.2, Supplementary Material online). UUG initiation is infrequent in fungal mitochondria (e.g., an isolated instance in *Mortierella*; GenBank #AY863211; supplementary fig. S2.2, Supplementary Material online). It does occur in bacteria where initiation of translation depends on a specific initiator methionine tRNA that recognizes AUG, GUG, and UUG, in descending order of efficiency. In the case of UUG initiation, only the second and third codon positions (U and G) are able to interact directly with the tRNA anticodon (i.e., the C and A of the CAU anticodon), accounting for the very low effectiveness of UUG as an initiation codon, despite stabilization and precise positioning of translation initiation codons by Shine–Dalgarno sequence motifs. These motifs occur at a defined distance upstream of initiation codons, yet in mitochondria, they are only known from the jakobid flagellate *Reclinomonas americana* (Lang et al. 1997). How then are UUG codons recognized in mitochondria of *G. rosea*?

A mechanism for more precise positioning of translation starts is known from mitochondria of monoblepharidalian chytrid fungi and the sea anemone *Metridium senile*, in which almost every protein-coding gene has a guanosine residue upstream of the predicted AUG or GUG start codons (Bullerwell et al. 2003). This conserved G residue converts the triplet into a more stable quadruplet initiation codon, correlating with the presence of an unorthodox cytosine residue at position 37 in the anticodon loop of the initiator tRNAs. This feature permits a more stable 4-bp

interaction between CAUC anticodons and quartet GAUG/GGUG codons. Inspection of *G. rosea* sequences does not confirm the presence of quadruplet initiation codons, yet its predicted initiator tRNA is highly unorthodox, with a G residue at position 32 that is otherwise always a pyrimidine (for a recent review on mitochondrial tRNA structure, see (Lang et al. 2011). In fact, in a 3D structure, position G-32 of the anticodon loop is in close vicinity to both the U-36 of the anticodon and the first U of UUG initiation codons. Interaction of G-32 with both U residues would thus stabilize the U–U pair. In other words, all three nucleotides of the UUG codon are capable of forming stable interactions with this unorthodox initiator tRNA, implying that translation initiation with UUG becomes as effective as with AUG: in accord with the high incidence of inferred UUG translation initiation codons in *G. rosea* mitochondria.

2.5.4 Robust Phylogenetic Association of Glomeromycota with Mortierellales

The Glomeromycota are part of zygomycetes, a phylogenetically heterogeneous (paraphyletic) taxon (James et al. 2006, Seif et al. 2005, Liu et al. 2009). According to a current proposal, Glomeromycota constitute an independent fungal phylum (Hibbett et al. 2007), yet phylogenomic analyses with nuclear and mitochondrial genes (e.g., (Seif, Cadieux et al. 2006, Liu, Leigh et al. 2009) do not support this view. When the complete mtDNA data of *G. irregulare* are included in the analysis, this fungus associates with *Mortierella* (Lang and Hijri 2009, Lee and Young 2009): an affinity that was also observed with a nuclear data set (Liu et al. 2009), although in both cases based on poor taxon sampling and without significant statistical support. Evidently, additional zygomycete nuclear or mitochondrial genome data are required to address this question and demonstrate beyond reasonable doubt that Glomeromycota merit the status of an independent phylum. When *G. rosea* data are included in a data set containing all

standard mtDNA-encoded proteins, Glomeromycota branch with *Mortierella* as noted above, but with better statistical support (PP 0.99; 66% jackknife value [JV]; fig. 2.2). In addition, *S. culisetae*, a rapidly evolving species belonging to Harpellales (grouping at a basal position with ML and the applied global model of protein evolution; data not shown), occupies a sister group position with Glomeromycota when PhyloBayes and the more realistic CAT model are used for phylogenetic inference (modeling site-wise heterogeneity of amino acid positions, one of the major sources of systematic phylogenetic error in global models) (Lartillot et al. 2004, Lartillot and Philippe 2007). However, the inclusion of *Smittium* in a monophyletic group including *Mortierella* and the two Glomeromycota draws only low statistical support (a posterior probability value of 0.68% and 43% JV), which we attribute to a combination of low phylogenetic signal (use of a single representative for Harpellales) and conflicting long-branch attraction (because rapidly evolving) toward a basal position in the tree. When removing *Smittium* from the data set, the support for *Mortierella* forming a monophyletic group with Glomeromycota increases substantially (PP 1.0; 93% JV)

In summary, our results further support the view that Glomeromycota and Mortierellales are sister taxa. This inference and the potential inclusion of Harpellales in the same monophyletic group remain to be tested with extended species sampling across ‘‘zygomycetes’’. In particular, representatives of Harpellales with more moderate evolutionary rates need to be identified and data from Mortierellales and Glomeromycota added. Contrary to our previously more pessimistic assessment (Lang and Hijri 2009), broad taxon sampling of mtDNAs alone may turn out to be sufficient to resolve phylogenetic relationships of Glomeromycota with confidence. Whatever the out-come, checking the consistency of mitochondrial phylogenies with phylogenomic analyses based on nuclear gene sequences will be important as a means of informing us about hidden phylogenetic artifacts.

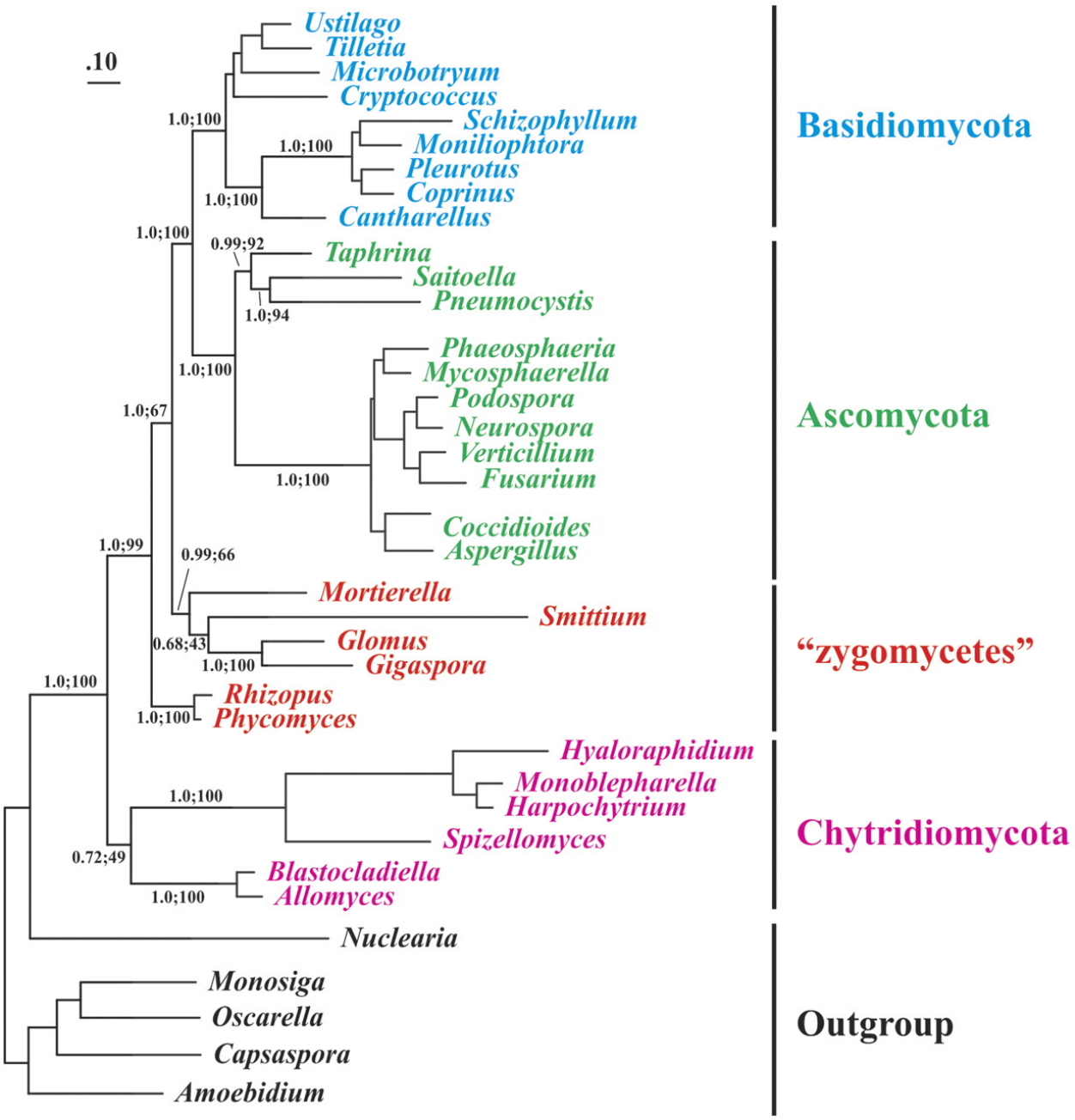


Figure 2.2. Phylogenetic positioning of Glomeromycota with mitochondrial protein data.

The tree shown was inferred with PhyloBayes, the CAT+Gamma model and six discrete categories (Lartillot and Philippe 2004) based on 13 concatenated proteins. The first number at branches indicates posterior probability values of a PhyloBayes analysis with four independent chains and the second number jackknife supports values. ML inference predicts a similar tree (not shown), except for *Smittium* and Blastocladiales.

2.5.5 Group I Intron–Mediated Trans-splicing of *G. rosea cox1*

The *G. rosea cox1* gene (encoding subunit 1 of respiratory chain complex IV) is broken up into two fragments that are located at a distance of ~30 kbp, encoded on the same strand of the mitochondrial genome. Formally, a *cox1* gene with a 30 kbp intron could be postulated, but this scenario is most unlikely for the following reasons. The two *cox1* fragments are separated by a total of 15 genes, including several tRNAs that are on the same strand. It is known that tRNAs are rapidly and efficiently processed from RNA precursors, so that a 30-kb RNA precursor would be rapidly fragmented within the large intervening region. In addition, the *cox1* break points correspond exactly to a position where *G. irregulare* carries a group I intron, favoring the hypothesis of group I intron–mediated trans-splicing. Examples of group I intron–mediated trans-splicing have been reported previously (Burger et al. 2009, Grewe et al. 2009, Pombert and Keeling 2010, Hecht et al. 2011) which in all three cases occur in the most intron-rich mitochondrial gene, *cox1*. The insertion points of trans-spliced group I introns vary (supplementary fig. S2.1, Supplementary Material online; table 2.2), with exceptions in the related species *Isoetes* and *Selaginella* and in *G. rosea* and *Trichoplax* where introns occupy the same (otherwise frequent) insertion point.

Table 2.2. Presence (•) or Absence (°) at Cognate Insertion Site of the Shared Introns between *Gigaspora* and *Glomus* in Representatives of Other Fungi.

Species	Accession	Introns								
Glomeromycota										
Glomus irregulare	FJ648425	rnli1939	nad5i417	nad5i717	°	cobi429	cox1i386	cox1i731	cox1i867	cox1i1108
Gigaspora rosea		•	•	•	nad5i954	•	•	±	•	•
Basidiomycota										
Cryptococcus neoformans	NC004336	°	°	°	°	°	°	°	°	°
Pleurotus ostreatus	EF204913	°	°	°	°	°	•	°	°	•
Ascomycota										
Pneumocystis carinii	GU133622	°	°	°	°	°	°	°	°	°
Podospora anserina	NC001329	°	°	°	°	°	°	°	•	•
Zygomycota										
Mortierella verticillata	AY863211	°	°	°	°	°	°	°	°	°
Rhizopus oryzae	AY863212	°	°	°	°	°	•	°	°	°
Chytridiomycota										
Monoblepharella	AY182007	°	°	•	°	°	°	°	°	•
A. macrogynus	NC001715	°	°	•	°	•	•	°	•	•

NOTE.—The representation ± indicates a partial group I intron secondary structure. The gray columns represent the *cox1* intron that is trans-spliced in *G. rosea*, and the *nad5* intron believed to be involved in the trans-splicing. Introns were named according to Dombrovska and Qiu (2004). Shared intron positions are based on the coding sequence of the gene in which they are inserted in the first Glomeromycota mitochondrial genome published (Lee and Young 2009). Unique introns to *G. rosea* are named according to their positions in the gene in which they were identified. Of the 13 group I introns in G. rosea and the 24 in *G. irregulare* 494, only seven are

shared at cognate insertion site. Note that the *G. rosea nad5*i954 intron that may complement a functional secondary structure of the trans-spliced *cox1* intron is not present in *G. irregulare* and in the other represented fungal species.

In the published examples, a complete canonical group I intron RNA secondary structure can be inferred by pairing the flanking portion of the discontinuous exons and is confirmed by use of group I intron prediction algorithms (Lang et al. 2007); Beck and Lang 2009). Yet in *G. rosea*, certain structural core elements appear to be either missing or to deviate significantly from established rules. As other group I introns in *G. rosea*, protein-coding genes do not display many unusual features, and as a catalytically conserved intron RNA core is essential for intron excision, we predict the use of an additional third helper RNA fragment in trans, similar to the situation of fragmented introns in *Chlamydomonas* (Goldschmidt-Clermont et al.1991) and *Oenothera* (Knoop et al. 1997). According to our assessment, the third intron in the *nad5* gene matches perfectly the two partial discontinuous intron sequences, completing a comprehensive intron RNA secondary structure (fig. 2.3A and B). To provide evidence that the mitochondrial *cox1* is not a pseudogene in *G. rosea* but indeed trans-spliced *in vivo*, we have conducted RT-PCR experiments, using primers located in the respective flanking exons. Sequencing of the resulting PCR product confirms that the *cox1* exons are accurately ligated *in vivo* (fig. 2.3C). Sequencing of the PCR product confirmed sequence identity with the respective *cox1* exons, without any sequence modification. In the placozoan animal example as mentioned above, trans-splicing combines with an additional RNA editing step (Burger et al. 2009). No PCR product was obtained with total genomic DNA, ruling out the possibility that a contiguous *cox1* gene exists in the nuclear genome.

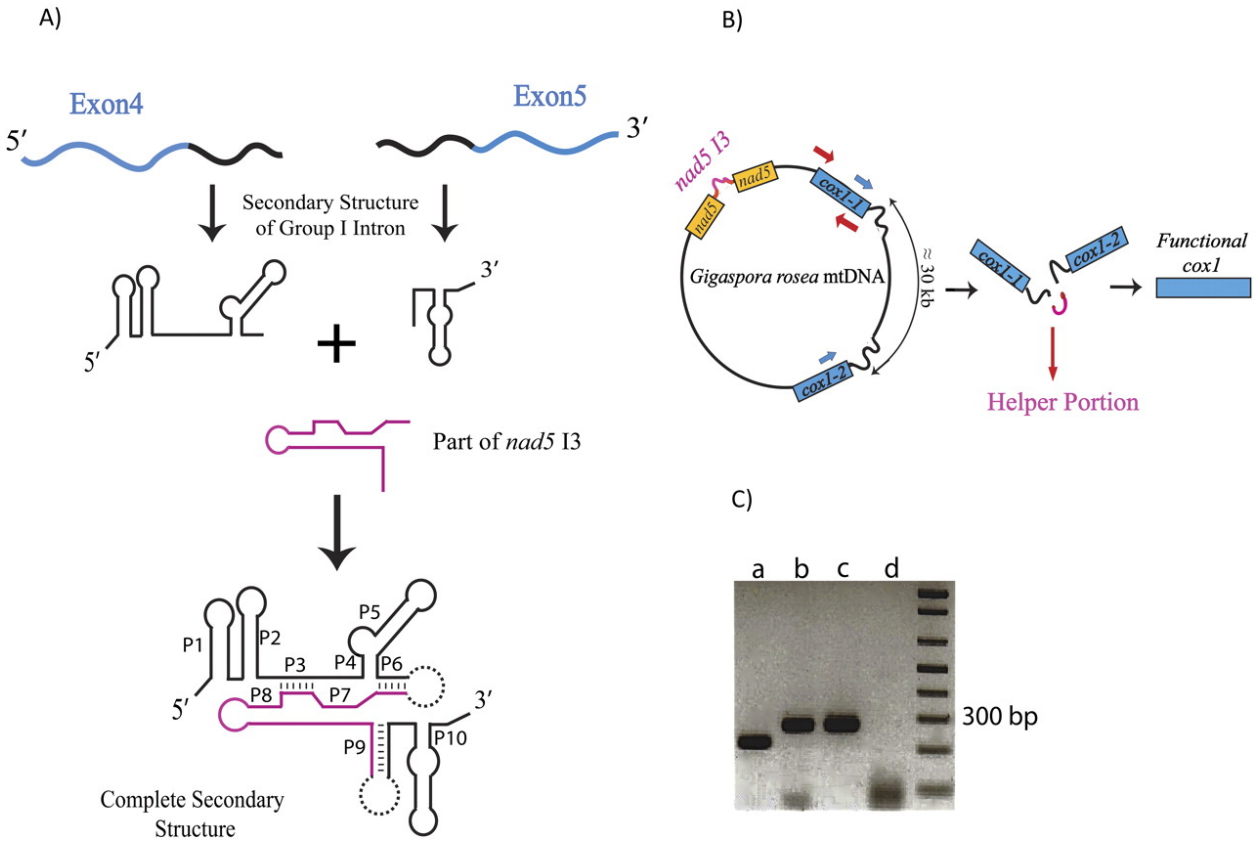


Figure 2.3. Model of group I intron–mediated trans-splicing in *G. rosea* and demonstration of mRNA trans-splicing by RT-PCR.

(*A*) Schematic view of the *trans*-spliced *cox1* group IB intron RNA secondary structure. The pink-colored RNA corresponds to the predicted *nad5* intron 3 helper sequence, and dashed loops indicate its interaction with the two intron fragments flanking *cox1* exons 4 and 5 (shown in blue). P1–P10 illustrate the conserved stems of the intron secondary structure according to Burke et al. (1987). Canonical Watson–Crick base pairing is shown by dashes. (*B*) Genomic view of RNAs involved in *cox*1 *trans*-splicing. Blue and red arrows indicate PCR and RT-PCR primers (see Materials and Methods). (*C*) Amplification products obtained with *G. rosea* cDNA and genomic DNA. a, RT-PCR amplification in the presence of total RNA (demonstration of *trans*-splicing and that mitochondrial *cox1* is not a pseudogene); b, RT-PCR amplification of an internal section of exon 4 (positive control); c, PCR amplification of total genomic DNA, with the same primers as in "b" (positive control); and d, PCR amplification of total genomic DNA, with the primers used in "a" (negative control, no presence of a contiguous gene in the nuclear genome). Further negative controls (see Materials and Methods) not shown.

2.5.6 Gene for the Small Subunit rRNA in Two Pieces

The other fragmented *Gigaspora* gene is *rns*, coding for the small subunit rRNA. In similar cases (including both the large and the small subunit rRNAs), trans-splicing has never been observed, and break points usually occur in variable loops of the rRNA secondary structure rather than in highly conserved regions as in protein-coding genes (Lang 1984). Secondary and tertiary interactions between the unligated rRNA fragments are sufficient for folding into a complete rRNA structure and formation of functional ribosomes. According to our rRNA secondary structure model, the breakpoint of *G. rosea* is in a variable loop (fig. 2.4). RT-PCR experiments did not produce DNA fragments of the predicted size and sequence (data not shown), and we explain a few incongruent RT-PCR sequences that we obtained by a low level of template-switching activity of RT. We have also been unsuccessful in folding the flanking regions into a bona fide intron RNA structures (either by inference with intron models that are part of the MFannot annotation or manually).

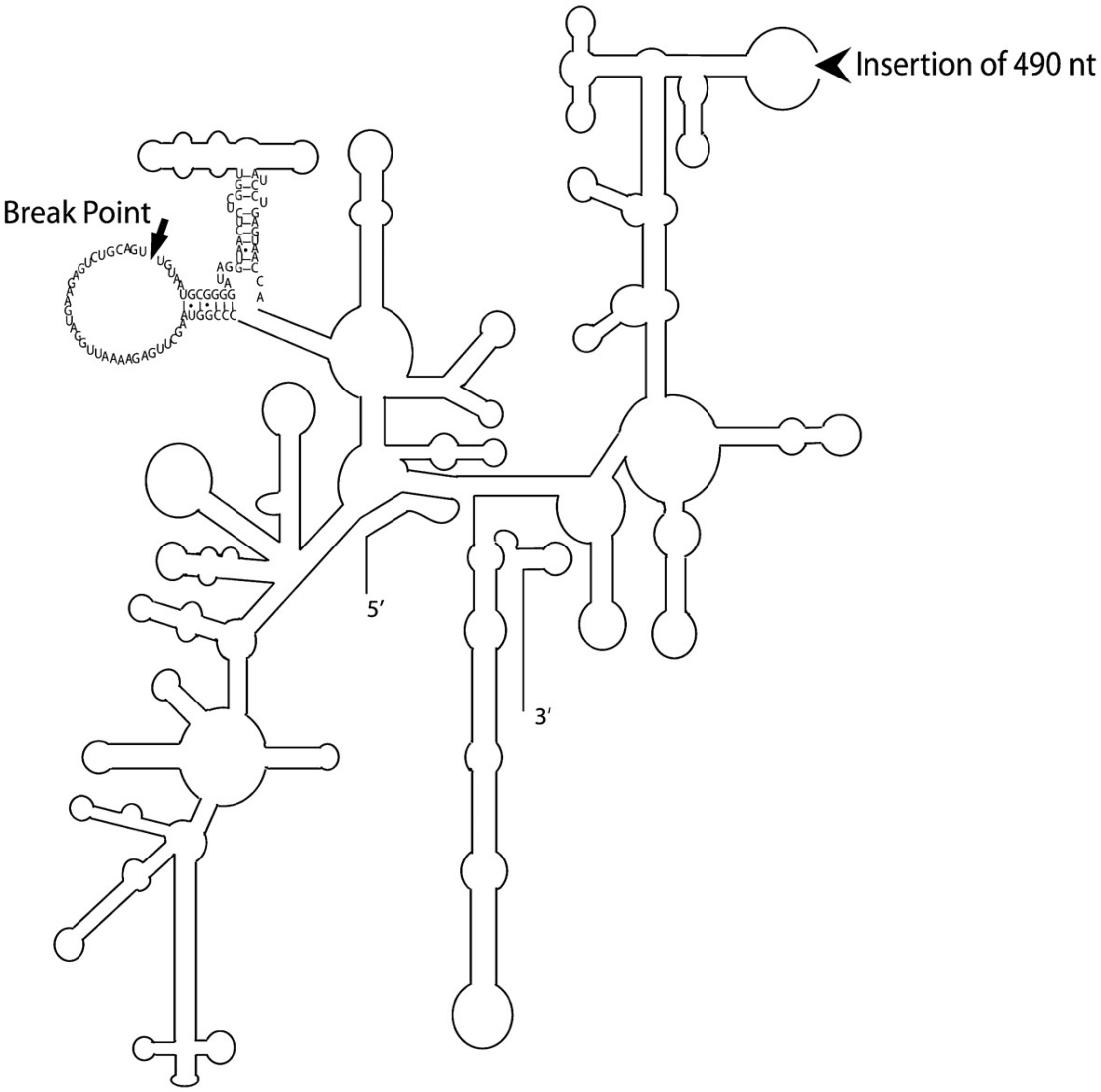


Figure 2.4. Secondary structure model of the fragmented *G. rosea* small subunit rRNA. Schematic *rns* structure in which only sequences close to the breakpoint are indicated. The 5′ and 3′ termini at the breakpoint are precisely inferred based on (significant) sequence similarity with the *Glomus irregulare* counterpart. Dashes indicate canonical Watson–Crick base pairings, and dots denote guanine–uracil pairings. The arrowhead points to a 490-nt (nonintron) insertion in a variable loop.

2.5.7 How Many True Introns in *rns* and *rnl*?

Modeling of rRNA structures and prediction of introns are problematic, without confirmation at the RNA sequence level. It is well known that mitochondrial rRNAs may carry inserts in variable regions of the structure that are not excised at the RNA level (e.g., (Lang et al. 1987)). In *G. rosea rns* and *rnl*, we find numerous inserts of this kind (figs. 2.2 and 2.4) that do not carry evidence for sound group I or group II intron structures. RT-PCR experiments confirm this view: the insert in *rns*_2 is not excised at the RNA level nor are five intron-size inserts in *rnl*. According to our results, RNA splicing occurs only for the two *rnl* introns that are predicted by RNAweasel (Beck and Lang 2009; one group II followed by a group I intron; fig. 2.2). It is noteworthy that one of the *rnl* inserts occurs precisely at a position where *G. irregulare* carries a true intron, as verified in that case by cDNA sequencing (Lee and Young 2009). Evidently, the nature and size of gene and genome insertions in glomeromycotan mtDNAs undergo the most rapid change, similar to what we observe with mitochondrial plasmid insertions.

2.6 Conclusion

The *Gigaspora* mitochondrial genome has turned out to be most unusual, with fragmented genes, trans-splicing, a ribosomal RNA gene in pieces, and an inflated genome size. The presence of substantial gene order differences between *Glomus* and *Gigaspora* is another intriguing observation, suggesting the capacity for nonhomologous recombination repair. The repeated insertion of plastid-like sequences into the mtDNA is in support of this view, as is a high but variable number of (mostly group I) introns. Group I introns propagate by a homing mechanism that relies on homologous recombination (Perrin et al. 1993, Chevalier et al. 2003) that is, we hypothesize that the biochemical machineries for both homologous and nonhomologous

recombination are present in glomeromycotan mitochondria. The effectiveness and mechanisms of recombination may be tested by combining, for instance, compatible *Glomus* species with differences in sequence or intron content via anastomosis and analysis for recombination products in the segregated offspring. Evidently, mtDNAs do segregate effectively in Glomeromycota, as any mtDNA so far sequenced has had a perfectly homogeneous sequence: in contrast to their nuclear genomes, which are multinucleate, most divergent in sequence, and seemingly impossible to assemble (Martin et al. 2008). The large differences among glomeromycotan mitochondrial genomes should motivate the future inclusion of other distant lineages, such as *Acaulospora* and *Scutellospora*, for further comparisons. These new sequences will at the same time serve in confirming the phylogenetic positioning proposed here of Glomeromycota within Fungi.

2.7 Supplementary Material

Supplementary figures S2.1 and S2.2 are available at Molecular Biology and Evolution online (http://www.mbe. oxfordjournals.org/).

2.8 Acknowledgments

We thank Dr Michael Gray (Halifax, NS) for critical reading of the manuscript, Natacha Beck for continued development of automated organelle genome annotation software (MFannot), and Dr Yves Piche´ for kindly providing *G. rosea*. We thank NSERC-CRD and Genome Quebec/Genome Canada for providing financial support to B.F.L. and M.H.

Presentation of article 2

During the past few years simultaneous with development of a low-cost 454 pyrosequencing method, first sequenced AMF mtDNA has been published (Lee and Young 2009). This publication was the pioneer of increasing number of sequenced AMF mtDNAs including more *Rhizophagus* spp. opening a new avenue through comparative mitogenomics analyses of representatives of AMF. Comparative mitogenomic analyses of *Rhizophagus irregularis* isolates would not only reveal AMF evolutionary history but also reveal potential suitable region for molecular marker development at low taxonomic level. MtDNA has been targeted in AMF to overcome complexities of their nuclear genome for identification purposes.

In this article, we compared the mitochondrial genomes of the newly sequenced AMF species *Glomus* sp. (Synonym: *Rhizophagus* sp.) with two isolates of the closely related *R. irregularis*. Along with a highly divergent intron insertion pattern, we found insertions of plasmid-related DNA polymerase and propagation of mobile open reading frame (mORFs) encoding endonucleases in *Rhizophagus* mtDNAs. Our findings have brought to light the first evidence of AMF interspecific exchange of mitochondrial coding sequences entailing formation of gene hybrids in *Rhizophagus* sp. *atp6*, *atp9* (coding for the subunit 6 and 9 of the ATP synthase complex), *cox2* (cytochrome C oxidase subunit 2) and *nad3* (NADH dehydrogenase subunit 3) genes.

This article was published in PLoS One. As a co-author, I contributed to mtDNA assembling and annotation. I explored the presence of hybrid genes in *G. diaphanum* isolate. D. Beaudet and B. Iffis continued these investigations and have done experiments on RNA. D. I have contributed to write the first draft of article and the final version was achieved by D. Beaudet and M. Hijri due to my maternity leave.

Chapter 3 - Rapid Mitochondrial Genome Evolution through Invasion of Mobile Elements in Two Closely Related Species of Arbuscular Mycorrhizal Fungi

Denis Beaudet, Maryam Nadimi, Bachir Iffis, Mohamed Hijri

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

Published in: PloS ONE, April 18, 2013. 8(4): e60768. doi: 10.1371/journal.pone.0060768

3.1 Abstract

Arbuscular mycorrhizal fungi (AMF) are common and important plant symbionts. They have coenocytic hyphae and form multinucleated spores. The nuclear genome of AMF is polymorphic and its organization is not well understood, which makes the development of reliable molecular markers challenging. In stark contrast, their mitochondrial genome (mtDNA) is homogeneous. To assess the intra- and inter-specific mitochondrial variability in closely related *Glomus* species, we performed 454 sequencing on total genomic DNA of *Glomus sp.* isolate DAOM-229456 and we compared its mtDNA with two *G. irregulare* isolates. We found that the mtDNA of *Glomus sp.* is homogeneous, identical in gene order and, with respect to the sequences of coding regions, almost identical t*o G. irregulare*. However, certain genomic regions vary substantially, due to insertions/deletions of elements such as introns, mitochondrial plasmid-like DNA polymerase genes and mobile open reading frames. We found no evidence of mitochondrial or cytoplasmic plasmids in *Glomus* species, and mobile ORFs in *Glomus* are responsible for the formation of four gene hybrids in *atp6*, *atp9*, *cox2*, and *nad*3, which are most probably the result of horizontal gene transfer and are expressed at the mRNA level. We found evidence for substantial sequence variation in defined regions of mtDNA, even among closely related isolates with otherwise identical coding gene sequences. This variation makes it possible to design reliable intra- and inter-specific markers.

2.2 Keywords

Arbuscular mycorrhizal fungi; Mitochondrial Genome; Comparative Mitochondrial Genomics; Plasmid related *dpo*; mobile endonuclease ORFs; Horizontal Gene Transfer.

3.2 Introduction

Arbuscular mycorrhizal fungi (AMF) are plant root-inhabiting obligate symbionts that form symbiotic associations with approximately 80% of plant species (Wang and Qiu 2006, Smith and Read 2008). This symbiosis helps plants to acquire nutrients and protects them from soil-borne pathogens (Azcón-Aguilar and Barea 1997, St-Arnaud and Vujanovik 2007) by inducing plant resistance (Datnoff et al. 1995, Cordier et al. 1998, Pozo et al. 2002, Ismail and Hijri 2012) or inhibiting pathogen growth (Ismail et al. 2011). In return, plants provide carbohydrates, which AMF cannot acquire from extracellular sources. They are an important component of soil microbial communities, as they are able to exchange their genetic material between compatible isolates through a process called anastomosis (Croll et al. 2009). The latter have been hypothesized to be an important factor in maintaining the genetic diversity found in Glomeromycota and to attenuate the effect of genetic drift within a population (Bever and Wang 2005, Corradi et al. 2007, Croll et al. 2009, Angelard and Sanders 2011, Colard et al. 2011). AMF are currently thought to reproduce clonally, based on the absence of a recognizable sexual stage (or apparatus). However, this hypothesis has been challenged by the identification of many orthologues of sexually-related genes (Halary et al. 2011, Sanders 2011, Tisserant et al. 2012), which suggests at least the presence of cryptic recombination. AMF spores and hyphae are multinucleated, but their true genetic organization is currently under debate (Kuhn et al. 2001, Pawlowska and Taylor 2004, Bever and Wang 2005, Hijri and Sanders 2005, Pawlowska and Taylor 2005). However, evidence strongly suggests that nuclei can be genetically divergent within an AMF individual. Thus, AMF are characterized by considerable within-isolate nuclear genetic diversity even at the expression level (Boon et al. 2010). The presence of such diversity in AMF individuals/populations (Boon et al. 2010, VanKuren et al. 2013), combined with a lack

of molecular data, have hindered the use of nuclear markers to assess questions on community structure, diversity and function. In contrast, AMF mitochondrial (mt) DNA is homogeneous within single isolates (Raab et al. 2005, Lee and Young 2009), making it a good target for marker development. Following this logic, the mitochondrial large subunit (LSU) rRNA gene has been explored for its usefulness as a marker (Raab et al. 2005, Borstler et al. 2008, Thiéry et al. 2010), although determining its specificity at the isolate level is still challenging for all AMF taxa aside from the model species *G. irregulare*.

Comparative AMF mitochondrial genomics has been proposed as an approach to open up new possibilities for development of strain-specific molecular markers (Lang and Hijri 2009, Lee and Young 2009, Formey et al. 2012) given that the type of mitochondrial marker necessary to establish specificity at different divergence levels may vary. This approach has been shown to be a powerful tool for the study of evolutionary relationships among lower fungi (Seif et al. 2005). Unfortunately, only three AMF mitochondrial genomes had been published until recently, including that of *Glomus intraradices* (Lee and Young 2009, Formey et al. 2012) (renamed to *G. irregulare* (Stockinger et al. 2009) and changed again recently to *Rhizophagus irregularis* based on an exhaustive molecular phylogeny of rRNA genes (Krüger et al. 2012); in the present paper, we will use the older nomenclature) as well as those of two distant AMF species, *Gigaspora rosea* and *Gigaspora margarita* (Nadimi et al. 2012, Pelin et al. 2012). Compared to *G. irregulare*, the Gigasporaceae genomes have an inflated mitochondrial genome size that is mainly the result of extended intergenic regions. These regions are not syntenic and both genomes harbor *cox1* and *rns* genes with exons encoded on different strands, whose products are joined at the RNA level through either trans-splicing events of group I introns, or basepairing. The mitochondrial protein sequences in the dataset were sufficient to confirm the phylogenetic relationship of AMF with *Mortierellales* as a sister group. This shows that a broader sampling of

AMF mtDNA can answer questions about the evolution of these ecologically important fungi. Formey et al. (2012) have recently sequenced four isolates of *G. irregulare* (Formey et al. 2012) and were able to develop isolate-specific markers using variable regions That were created by the insertion of mobile elements.

Those elements, including linear or circular plasmids and mobile ORF encoding endonucleases (mORFs), are present in a broad range of fungal mitochondrial genomes (For review see (Hausner 2012)). Plasmids are autonomously-replicating circular or linear extrachromosomal DNA molecules. They are found in three broad types: circular plasmids encoding a DNA polymerase gene (*dpo)* (Griffiths and Yang 1995), linear plasmids with terminal inverted repeats encoding either a *dpo* or *rpo* (RNA polymerase) gene or both (Klassen and Meinhardt 2007) and retroplasmids, which usually encode a reverse transcriptase (Kennel and Cohen 2004). Free linear or circular plasmids encoding *dpo* can be present in the mitochondria of fungi (Griffiths and Yang 1995) and plants (Brown and Zhang 1995). Segments have been shown to integrate within the mtDNA of fungi (Bertrand and Griffiths 1989, Ferandon et al. 2008), but plasmid-related *dpo* insertions tend to fragment, shorten (since they are not selected for) and eventually disappear from mitochondrial genomes. Plasmid-related *dpo* insertions have been reported in the AMF *Gigaspora rosea*, but are virtually absent from the closely related paraphyletic zygomycetes. The mobility of mORFs, elements that thrive in *Glomus*, is mediated by the site-specific DNA endonuclease they encode. This endonuclease cleaves ORF-less alleles by creating a double-strand break in DNA and initiates the insertion and fusion of the mobile element. The same process, called *intron homing*, has been proposed for group I introns (Dujon 1980). Several lines of phylogenetic evidence support the hypothesis of the evolutionary-independent ancestral origins of mORFs (Bell-Pedersen et al. 1990). These highly mobile elements have the ability to carry group I introns (Dalgaard et al. 1993), intergenic sequences

(Sharma et al. 1992), and coding sequences (Eddy 1992). The first reported case of mitochondrial gene transfer caused by those elements was a mORF-mediated insertion of a foreign *atp6* carboxy-terminal in the blastocladiomycete *Allomyces* macrogynus (Paquin et al. 1994).

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

3.3 Materials and Methods

3.3.1 Fungal material

Spores and mycelium of *Glomus sp.* (DAOM 229456*)* and *G. irregulare* (DAOM 197198) were cultivated *in vitro* on a minimal (M) medium with carrot roots transformed with *Agrobacterium rhizogenes*, as described in the literature (Becard and Fortin 1988). The medium was liquefied using a 0.82 mM sodium citrate and 0.18 mM citric acid extraction buffer solution. The resulting fungal material was further purified by hand under a binocular microscope, to remove root fragments.

3.3.2 DNA extraction

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

3.3.3 RNA extraction

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

3.3.4 cDNA synthesis

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

ribonuclease (Promega, Canada) was added to the cDNA and incubated at 37°C for 20 min. The resulting cDNA was stored at −20°C until use.

3.3.5 Polymerase chain reaction (PCR)

The proposed intergenic markers to discriminate between *G. irregulare* DAOM197198 and *Glomus sp.* were tested by PCR using the KAPA2G Robust Hotstart ReadyMix PCR kit (KapaBiosystems, Canada). The specific primers used were respectively *rnl-cox2*_197198_spec_F (5′-AAAGGAATTACATCGATTTA-3′), *rnl-cox2*_197198_spec_R (5′-ACAAGAAGGTTTGCATCGCTA-3′), *nad6-cox3*_dia_spec_F (5′-CCACTAGTTAAGCTACCCTCTA-3′) and *nad6-cox3*_dia_spec_R (5′-AATCATACCGTGTGAAAGCAAG-3′). The variable length primers were *rnl-cox2*_197198_size_F (5′-TAGGGATCAGTACTTTAGCCAT -3′), *rnl-cox2*_197198_size_R (5′-TCCTTACGGTATGAATGGTAAG-3′), *rnl-cox2*_dia_size_F (5′-AGACTTCTTCAGTTCCACAATCA-3′) and *rnl-cox2*_dia_size_R (5′-ATGGCTAAAGTACTGATCCCTAC-3′) (Table 3.1).. For 40 µl of PCR reaction volume, 12 µl of water, 20 µl of 2× PCR buffer, 3.5 µl of (5 µM) forward and reverse primers, and 1 µl of DNA were added. Cycling parameters were 94°C/3 min, followed by 38 cycles of: 94°C/30 sec, 54°C/25 sec, 72°C/45 sec and a final elongation at 72°C. PCR products were separated by electrophoresis in a 1.5% (w/v) agarose gel and visualized with GelRed under UV light.

3.3.6 Reverse transcriptase – polymerase chain reaction (RT-PCR)

Our objective with regard to PCR reactions on cDNA was to assess which regions of the

gene hybrid reported in *atp6*, *atp9*, *cox2* and *nad3* were expressed at the mRNA level. For each of the four hybrids, a forward primer designed in the conserved 'core' structure of the gene (*atp6*_core_F: 5′-AGAGCAGTTTGAGATTGTTAAG-3′, *atp9*_core_F: 5′-CTGGAGTAGGAGTAGGGATAGT-3′, *cox2*_core_F: 5′-CATGGCAATTAGGATTTCAAGA-3′ and *nad3*_core_F: 5′-TCGTTCCTTTGTTCGTGCTA-3′) was used in combination with three reverse primers designed respectively in the inserted C-terminal (*atp6*_insert_R1: 5′-AGCCTGAATAAGTGCAACAC-3′, *atp9*_insert_R1: 5′- GTAAGAAAGCCATCATGAGACA-3′, *cox2*_insert_R1: 5′-TGAGAAGAAAGCCATAACAAGT-3′ and *nad3*_insert_R1: 5′-AGAAGTATGAAAACCATAGCAATC-3′), the mobile ORF (*atp6*_mORF_R2: 5′-AGTCTTCGAATATACTGGCAG-3′, *atp9*_mORF_R2: 5′-TGTCGAGTCTCCAAAGTATGT-3′, *cox2*_mORF_R2: 5′-ACTGAATTCCTGTGTTTCGATCT-3′ and *nad3*_mORF_R2: 5′-TGACGAATGGTTAGACGATGT-3′) and the native C*-terminal portion of the corresponding gene (*atp6*_native_R3: 5′-CGTACCGTCGTAACAAGTAGA-3′, *atp9*_native_R3: 5′-CCATCATTAAGGCGAATAGA-3′, *cox2*_native_R3: 5′-CTAACAAACTCCCGACTATTACCT-3′ and *nad3*_native_R3: 5′-AGAATGAAGACCATTGCAAC-3′). To verify that there was no residual mitochondrial DNA in the cDNA, the primers Ctrl_positive_*nad5*exon4_689F (5′-ACCATTCTGTTATGTTCTAATGT-3′) and Ctrl_positive_*nad5*exon4_689R (5′-GTCTGACTTAGCAGGTTAGTTAAG-3′) were designed in *nad5* exon 4 and used as a positive control on cDNA and negative control on RNA (Table 3.1). The RT-PCR reactions were carried out using the KAPA2G Robust Hotstart ReadyMix PCR kit (KapaBiosystems, Canada) as described above in the PCR section.

Table 3.1. List of primers for PCR and RT-PCR.

	specific primers	Primers
PCR	*rnl-cox2*_197198_spec_F	5′-AAAGGAATTACATCGATTTA-3′
	*rnl-cox2*_197198_spec_R	5′-ACAAGAAGGTTTGCATCGCTA-3′
	*nad6-cox3*_dia_spec_F	5′-CCACTAGTTAAGCTACCCTCTA-3′
	*nad6-cox3*_dia_spec_R	5′-AATCATACCGTGTGAAAGCAAG-3′
	variable length primers	Primers
	*rnl-cox2*_197198_size_F	5′-TAGGGATCAGTACTTTAGCCAT -3′
	*rnl-cox2*_197198_size_R	5′-TCCTTACGGTATGAATGGTAAG-3′
	*rnl-cox2*_dia_size_F	5′-AGACTTCTTCAGTTCCACAATCA-3′
	*rnl-cox2*_dia_size_R	5′-ATGGCTAAAGTACTGATCCCTAC-3′
RT-PCR	*atp6*_core_F	5′-AGAGCAGTTTGAGATTGTTAAG-3′
	*atp9*_core_F	5′-CTGGAGTAGGAGTAGGGATAGT-3′
	*cox2*_core_F	5′-CATGGCAATTAGGATTTCAAGA-3′
	*nad3*_core_F	5′-TCGTTCCTTTGTTCGTGCTA-3′)
	*atp6*_insert_R1	5′-AGCCTGAATAAGTGCAACAC-3′
	*atp9*_insert_R1	5′- GTAAGAAAGCCATCATGAGACA-3′
	*cox2*_insert_R1	5′-TGAGAAGAAAGCCATAACAAGT-3′
	*nad3*_insert_R1	5′-AGAAGTATGAAAACCATAGCAATC-3′
	*atp6*_mORF_R2	5′-AGTCTTCGAATATACTGGCAG-3′
	*atp9*_mORF_R2	5′-TGTCGAGTCTCCAAAGTATGT-3′
	*cox2*_mORF_R2	5′-ACTGAATTCCTGTGTTTCGATCT-3′
	*nad3*_mORF_R2	5′-TGACGAATGGTTAGACGATGT-3′
	*atp6*_native_R3	5′-CGTACCGTCGTAACAAGTAGA-3′
	*atp9*_native_R3	5′-CCATCATTAAGGCGAATAGA-3′
	*cox2*_native_R3	5′-CTAACAAACTCCCGACTATTACCT-3′
	*nad3*_native_R3	5′-AGAATGAAGACCATTGCAAC-3′
	Ctrl_positive_*nad5*exon4_689F	5′-ACCATTCTGTTATGTTCTAATGT-3′
	Ctrl_positive_*nad5*exon4_689R	5′-GTCTGACTTAGCAGGTTAGTTAAG-3′

3.3.7 Cloning

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

3.3.8 Sequencing, assembly and gene annotation

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

3.3.9 Phylogenetic analysis

For each gene of interest (*atp6, atp9, cox2* and *nad3*) in 12 AMF species, the dataset contains the corresponding C*-terminal for: *Glomus sp.* DAOM-229456, *G. irregulare* isolate 494, *G. irregulare* DAOM-197198, *G. irregulare* DAOM-240415, *G. irregulare* DAOM-234179, *G. irregulare* DAOM-234328, *G. irregulare* DAOM-213198, *Glomus sp.* DAOM-240422, *G. fasciculatum* DAOM-240159, *G. aggregatum* DAOM-240163, *G. cerebriforme* DAOM-227022, *Gigaspora rosea* DAOM-194757 (accession number JQ693396) and 3 selected fungal representatives: *Mortierella verticillata* (accession number AY863211), *Smittium culisetae* (accession number AY863213) and *Rhizopus oryzae* (accession number AY863212). The sequences were deposited in databases under the accession numbers: JX074786-JX074817. The reference phylogeny was constructed using the concatenated 'core' sequence (without the C*-

terminal portion used previously) of the same four genes. The DNA sequence alignments and the inference of maximum likelihood trees using GTR+G (with five distinct gamma categories) were performed using the integrated program MEGA version 5 (Tamura et al. 2011). Bootstrap resampling (1000 replicates) was carried out to quantify the relative support for each branch of the trees. Bayesian analysis were done using MrBayes version 3.2 using the GTR+G model (with five distinct gamma categories), four independant chains, one million cycles, tree sampling every 100 generations and a burn-in value of 25%.

3.4 Results and Discussion

3.4.1 *Glomus sp.* genome organization and structure

The complete sequence of the *Glomus sp.* 229456 mt genome was a double-stranded circular DNA molecule, exempt of polymorphism, with a size of 87,763 bp. The annotated sequence of *Glomus sp.* was deposited in GenBank under the accession number JX065416. Its mtDNA harbors the typical set of 41 mitochondrial genes found in other AMF (two rRNAs, 14 protein coding genes (PCGs) and 25 tRNAs). The PCGs include three ATP synthetase (*atp*), one cytochrome b (*cob*), three cytochrome C oxydase (*cox*) and seven NADH dehydrogenase (*nad*) genes. Also, 19 ORFs and 31 introns are inserted in this newly sequenced mt genome (Figure 3.1).

3.4.2 Comparative view of three *Glomus* mtDNAs

The gene content in *Glomus sp.* and *G. irregulare* mitochondrial genomes is similar to that found in zygomycetes, except for *rps3* and *rnpB*. The mtDNAs of both AMF species have the same gene order, and all genes are transcribed from one strand with very similar coding regions

except for the insertion of mobile ORF elements (mORFs) in the *atp6, atp9*, *cox2* and *nad3* genes of *Glomus sp.* (Figure 3.2).

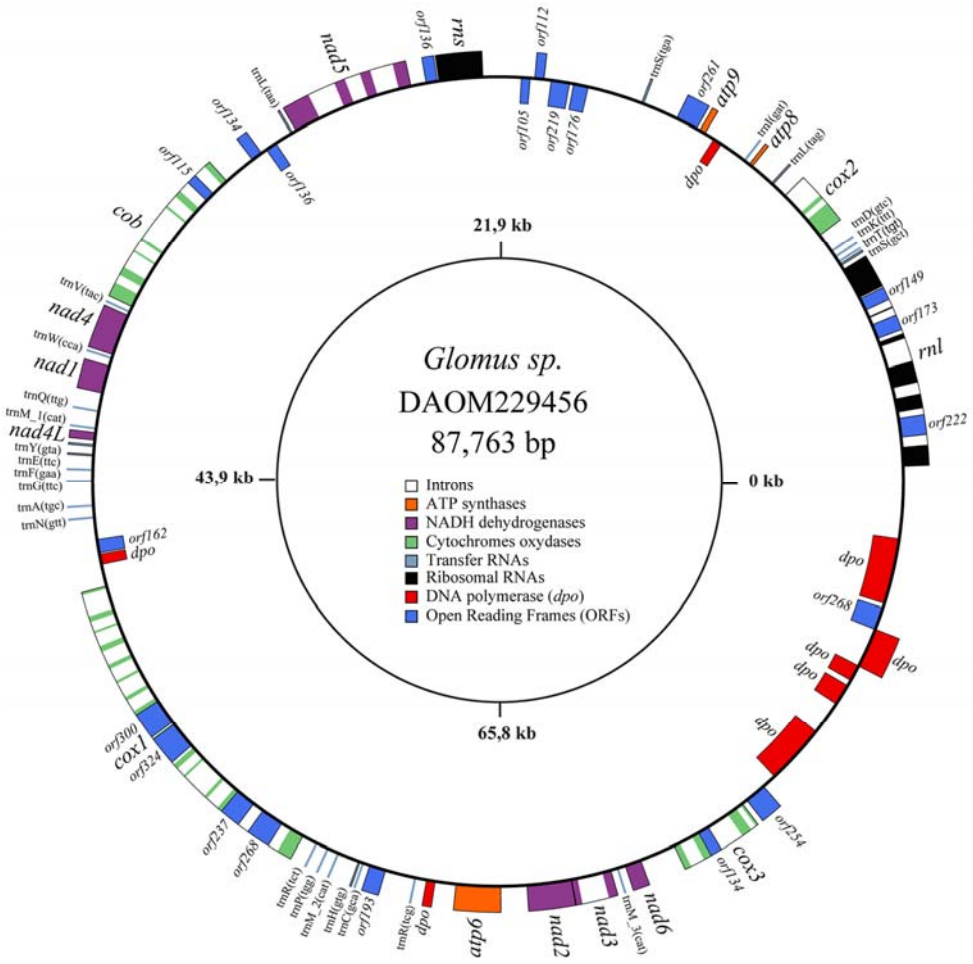


Figure 3.1. The *Glomus* sp. 229456 mitochondrial genome circular-map was opened upstream of *rnl*.

Genes on the outer and inner circumference are transcribed in a clockwise and counterclockwise direction, respectively. Gene and corresponding product names are *atp6, 8, 9,* ATP synthase subunit 6; *cob*, apocytochrome b; *cox1–3*, cytochrome c oxidase subunits; *nad1–4, 4L, 5–6*, NADH dehydrogenase subunits; *rnl, rns*, large and small subunit rRNAs; A–W, tRNAs, the letter corresponding to the amino acid specified by the particular tRNA followed by their anticodon. Open reading frames smaller than 100 amino acids are not shown.

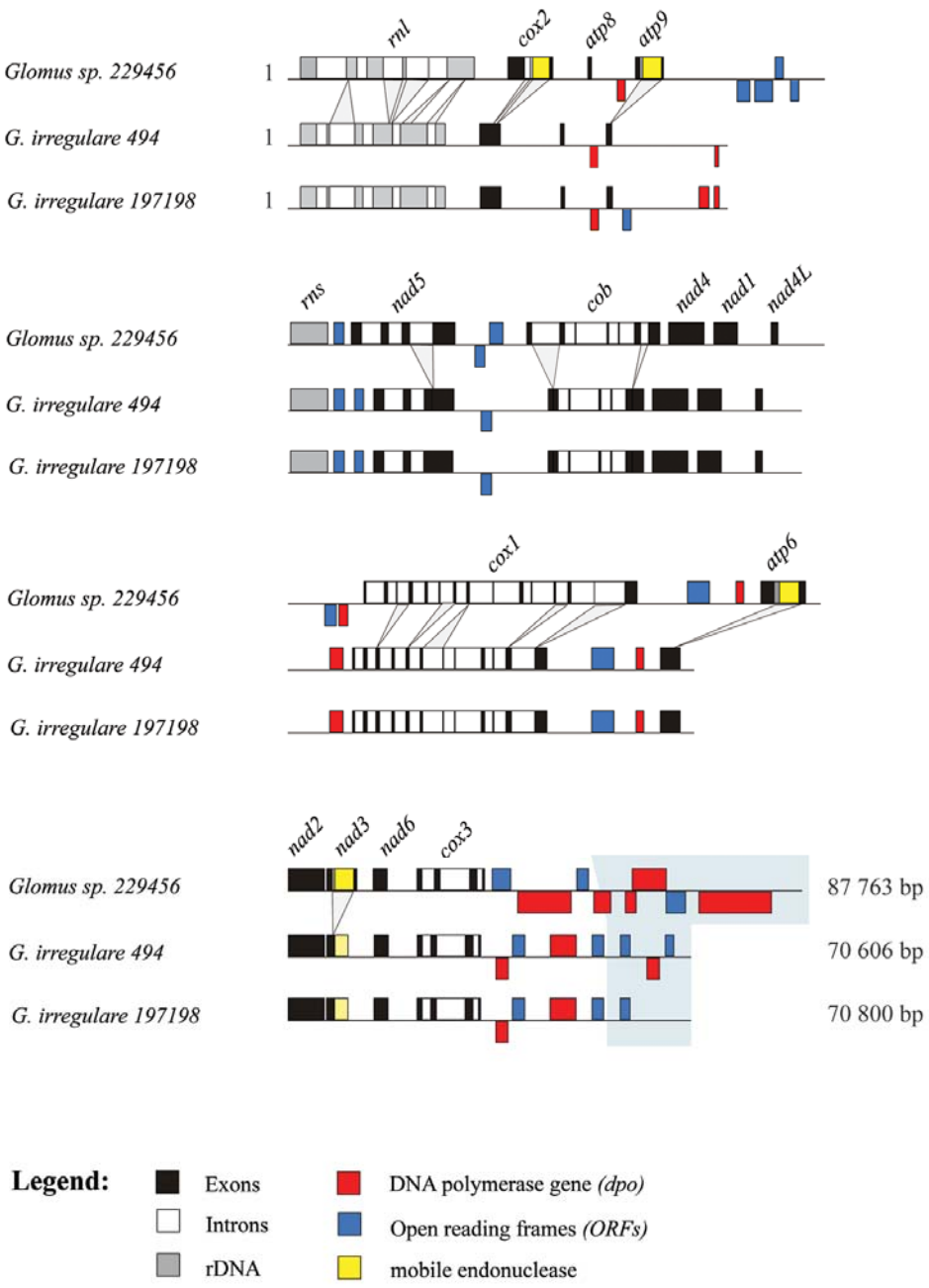


Figure 3.2. Comparative view of the three mitochondrial genomes linear map where the exons (black), introns (white), rDNA (gray), *dpo* plasmid insertions (red), ORFs (blue) and mobile endonuclease (yellow) are represented.

Divergence in intron insertion pattern is indicated by projections. A hyper-variable region in the *cox3-rnl* intergene is boxed in grayscale.

However, there are many differences in the number of introns, some of which carry more substantial sequence differences than do the coding sequences. The differences in the presence of introns and mORFs explain the inflated genome size of 87,763 bp in *Glomus sp.*, as compared to 70,800 bp in *G. irregulare* 197198 (Table 3.2). *Glomus sp*. *cox1* intron 8 is the homolog of an intron inserted at the same position in *Rhizopus oryzae* and angiosperms (with 76 and 79% of sequence identity, respectively) (Lang and Hijri 2009). *G. irregulare cox1* intron 7 is also inserted at the same position, but has an eroded ORF encoding the homing endonuclease gene, and thus also shares identity with the intron RNA secondary structure of *R. oryzae* and plants. The plant *cox1* intron was thought to have been acquired from a fungal donor, due to the proximity of its clade to that of fungi rather than to the non-vascular plant *Marchantia*. Knowing the extent to which the intron has spread in angiosperms (Cho et al. 1998, Sanchez-Puerta et al. 2008), it would be interesting to see whether such an invasion has also occurred within the Glomeromycota phylum.

Table 3.2. Gene and intron content in AMF and selected fungal mtDNAs.

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

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

Further, intergenic regions differ substantially in sequence: some are identical while others show signs of very fast, substantial changes including point mutations, insertions, deletions and inversions (Figures 3.2 and 3.3). Most of these differences occur in the *cox3-rnl* intergene, a large hyper-variable region that has been invaded by *dpo* fragments. The variations observed in intergenic regions provide an opportunity to develop species- specific molecular markers as shown in Figure 3.3, and even isolate-specific markers or methods allowing reliable identification and/or quantification of these fungi.

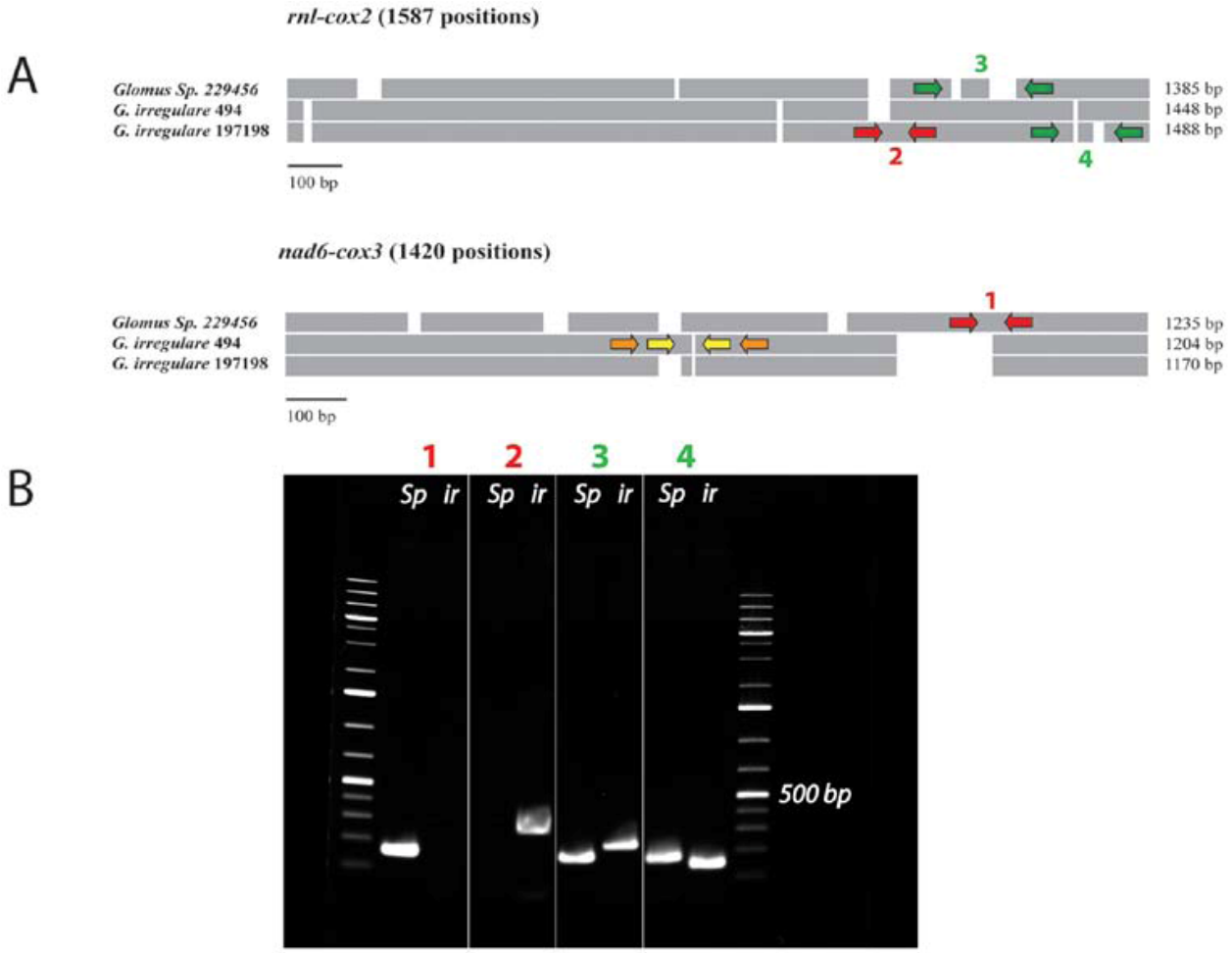


Figure 3.3. Schematic alignment representation of two mitochondrial intergenic regions (*rnl-cox2* and *cox3-nad6*) showing the presence of numerous insertions and deletions (indels)

A) The red arrows indicate the approximate position of the PCR primers that yield strain-specific markers, while the green arrows indicate the position of PCR primers that produce a size-specific marker. The yellow and orange arrows indicate potential regions to design, respectively, specific and size-specific markers in *G. irregulare* 494. B) Agarose gel electrophoresis figure showing the PCR results of the proposed markers on *Glomus sp.* 229456 (*Gs*) and *G. irregulare* DAOM197198 (*Gi*) respectively for each marker. Marker 1 shows the *Glomus sp.* specific amplification (156 bp), while marker 2 shows the *G. irregulare* 197198 specific marker (263 bp). The size-specific marker 3 yield a length of 160 bp for *Glomus sp.* and 226 bp for *G. irregulare* 197198. Finally, the size-specific marker 4 yield a length of 159 bp for *Glomus sp.* and 131 bp for *G. irregulare* 197198.

Lack of efficient and powerful molecular markers for AMF identification and quantification constitutes a major problem that limits the analysis of population genetics and field studies in AMF. Mitochondrial DNA is homogeneous within the AMF individuals studied to date, but evidence of genetic polymorphism between *G. irregulare* isolates has been observed in intergenic regions. They harbor highly conserved genes as well as highly variable regions, which promises to facilitate AMF barcoding at different taxonomic levels, an analysis that is currently challenging to carry out using nuclear genes. Hyper-variable intergenic regions with eroded *dpo* insertions and indels in intergenic regions constitute useful mitochondrial areas on which to focus attention in order to develop suitable markers for discriminating isolates of the same species. Intron insertion pattern variations, genome reorganizations (such as gene shuffling) and coding region divergences will make it possible to distinguish between different AMF species, genera and families. Our 454 pyrosequencing data and direct PCR sequencing showed that *G. irregulare* DAOM197198 and *Glomus sp.* mtDNAs are homogeneous, meaning that all the mitochondrial genomes in a given isolate are essentially identical, in stark contrast to the nuclear genomes. Our results confirm the previous report by Lee et al. (2009) suggesting homoplasmy in the first completed Glomeromycota mitochondrial genome of the AMF *G. intraradices* (*G. irregulare* isolate 494). A rapid and effective mitochondrial segregation mechanism was suggested to explain those findings. It was previously demonstrated that isolates of the same species can exchange nuclear material through anastomosis (Croll et al. 2009), but exchange of divergent mitochondrial haplotypes has yet to be shown. This leads us to question whether polymorphism does indeed occur through anastomosis, and for how many generations mitochondrial heteroplasmy is maintained.

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

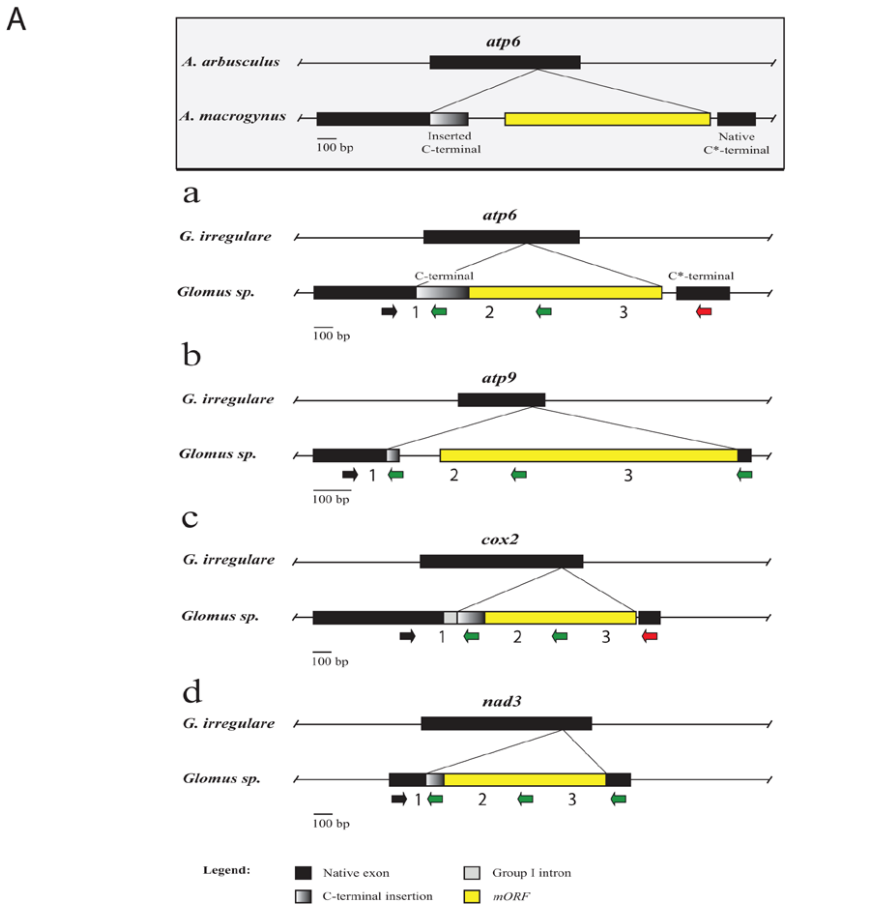
Plasmid-related DNA polymerase genes are found in mobile mitochondrial plasmids that occur either as free linear or circular DNAs, and have been shown to also insert into mtDNA (for review see (Hausner 2012). One striking feature in the comparison of the two closely related *Glomus* species is the presence of numerous *dpo* insertions in the intergenic regions of their mtDNA (Figure 3.2). All three *Glomus* mtDNAs contain a large number of *dpo* fragments, most of which are substantially divergent in sequence and therefore are most likely the result of independent plasmid insertion events. Even the two *G. irregulare* (isolates 494 and DAOM197198), otherwise almost identical in sequence, differ in *dpo* sequence, which supports the interpretation that *dpo* insertion occurs repeatedly and frequently through evolutionary time. A *bona fide* and complete *dpo* gene is present in *Glomus sp*., and its sequence is different from those in *G. irregulare* isolates. Because of its complete length, it most likely results from a recent insertion event. There is no evidence that *dpo* is functional when inserted in mtDNA. As in numerous other cases, *dpo* coding regions are fragmented in *Glomus* and occur on both strands, representing a good indicator of a genomic region experiencing little if any selective evolutionary constraints. The source of the *dpo* insertions in *Glomus* mtDNA remains elusive. We did not find any free mitochondrial plasmids in our *Glomus sp.* and *G. irregulare* isolate DAOM197198 shotgun data (combining nuclear and mitochondrial DNAs), as we did for *Gigaspora rosea*, where a 3582 bp contig with high sequence coverage was found (Nadimi et al. 2012). However, since the *Glomus* strains used in this study come from aseptic *in vitro* cultures, and even though the *G. rosea* fungal material was extracted from *in vivo* greenhouse pot cultures, we cannot rule out the possibility that an environmental vector is the source of *dpo* plasmids and is responsible for their propagation in *G. rosea*. Interestingly, *dpo* plasmids have been found to occur in numerous plants, notably in *Daucus carota* (Robison and Wolyn 2005) which is used as a host

plant for AMF cultures *in vitro*. The obligate biotrophic dependence of AMF on plants could be one of the reasons that *dpo* insertions are most abundant in *Glomus* mtDNAs yet virtually absent in mitochondrial sequences of the Blastocladiomycota (except for a single 100 amino acid long fragment occurrence in *Smittium culisetae* mtDNA), which is the closest phylogenetic group to the Glomeromycota. Mobile element insertions have been shown to trigger genomic rearrangements such as gene shuffling through homologous recombination (Brügger et al. 2004) and even genome linearization (Biessmann et al. 1992, Fricova, Valach et al. 2010, Hausner 2012). Whenever sequence repeats occur, more than one genome conformation may exist, but we have no evidence that this happens in *Glomus* mtDNA. It would be interesting to examine whether numerous recent *dpo* insertions with high sequence similarity might act as genomic repetitions and give rise to genome reorganization in closely related AMF species. Integrated plasmid segments within mitochondrial genomes, even though they are neutral or cryptic, could promote genomic rearrangements.

3.4.4 Mobile ORF elements (mORFs) in *Glomus*

Although most ORF-encoding endonuclease genes are inserted in introns where they have been shown to play a role in propagation, they can also be present in genes in which their evolutionary impact is less obvious. We identified numerous mORFs encoding endonuclease genes unique to *Glomus sp.* isolate DAOM-229456 mtDNA. When we annotated the sequences of the *atp6*, *atp9*, *cox2* and *nad3* genes, we observed that they all have a peculiar organization. Indeed, these genes harbor a carboxy-terminal duplication (C*-terminal) that was found downstream of a mORF insertion. For example, in the *atp6* gene, the duplicated portion of the C-terminal was found about 1000 bp downstream, following an inserted LAGLIDADG

endonuclease ORF. When we compared the DNA sequence of the C*-terminal portion with the corresponding sequence of *G. irregulare* isolate 494, a close relative to *Glomus sp.*, we found a 91.2% nucleotide identity. In contrast, the comparison between the *Glomus sp. atp6* duplicated carboxy-terminals (C-terminal and C*-terminal) showed a low sequence identity of 63.5%. Interestingly, comparison of the *Glomus sp*. C*-terminal amino acid sequences with the corresponding portion in *G. irregulare* showed 100% identity, indicating that the mutations observed in DNA are all synonymous. However, the comparison of the amino acid sequences of *Glomus sp*. *atp6* carboxy-terminals showed 91% identity. Surprisingly, when we designed a forward primer in the upstream sequence (5′ gene portion) and two reverse primers in the C-terminal and C*-terminal respectively, we found that the C-terminal is transcribed with the upstream sequence resulting in a putative hybrid transcript while the C*-terminal was not expressed into mRNA. Thus we hypothesized that the C-terminal portion could have been acquired from a donor through horizontal gene transfer (HGT). We also observed similar organization in *atp9*, *cox2* and *nad3* genes of *Glomus sp.* where the carboxy-terminal portion (C*-terminal) was replaced partially or completely by one carried by a mORF (C-terminal) encoding a LAGLIDADG endonuclease (except in *atp9*, a GIY-YIG family endonuclease) (Figure 3.4A: a, b, c and d). In *atp6*, the insert lacks a stop codon and the ORF is in phase with the native gene. In *atp9*, the insert, along with the mORF, completely replaces the native 3′ end, while in *cox2* and *nad3* only a portion of the carboxy-terminal is replaced (Table 3.3). The resulting gene hybrids are expressed at the mRNA level in all four cases as shown in Figure 3.4B. After sequencing of the cDNA bands, we found that the mORF and the inserted C-terminal are integral parts of the transcript in all four genes. However, in *atp6* and *cox2*, the native C*-terminal was not expressed into mRNA.



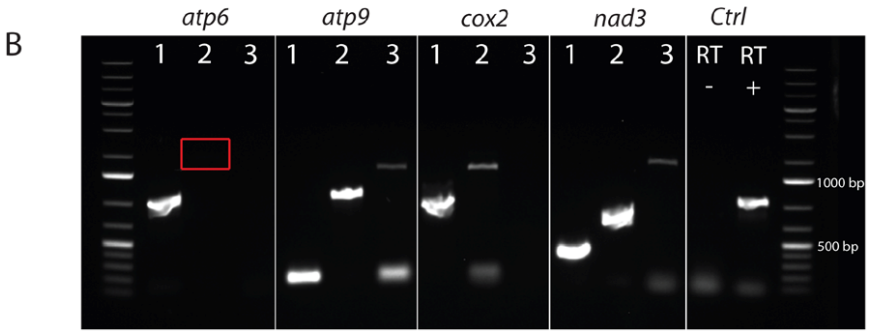


Figure 3.4. Comparison of gene hybrids found in *atp6*, *atp9*, *cox2* and nad3.

A) The *atp6* gene hybrid reported for *Allomyces macrogynus* (grayscale, boxed) is used as a reference in a comparison of the most similar *atp6* (a), *atp9* (b), *cox2*(c) and *nad3* (d) genes in *Glomus sp.* mtDNA. Each occurrence is put in perspective with the gene of a close relative (either *Allomyces arbusculus* or *G. irregulare*) in order to show the insertion point of the foreign element with the projections. Exons are in black, while the inserted foreign C-terminal is shaded in gray. The mobile endonuclease element is in yellow. For each gene, the black arrow indicates

the position of the forward primer used in the downstream RT-PCR experiment in combination with three different reverse primers. The green arrows indicate expression at the RNA level of the corresponding portion of the gene, while the red arrows indicate a negative amplification. B) Agarose gel electrophoresis figure showing the RT-PCR results. For each gene hybrid, the expression at the RNA level was tested using a forward primer in the conserved gene core and a reverse primer respectively in the inserted C-terminal (1), the mobile endonuclease (2) and the native C*-terminal portion (3). Primers in *nad5* exon 4 were used as a positive control on cDNA (RT +) and negative control on RNA (RT −). The expected size of the amplified fragments was: *atp6* inserted C-terminal (684 bp), *atp9* inserted C-terminal (149 bp), *atp9* mORF (717 bp), *atp9* native C*-terminal (1085 bp), *cox2* inserted C-terminal (938 bp), *cox2* mORF (1291 bp), *nad3*inserted C-terminal (261 bp), *nad3* mORF (597 bp), *nad3* native C*-terminal (1183 bp) and the positive control in *nad5* exon 4 had an expected amplicon size of 689 bp. The red box indicates a faint band that is present on the gel.

Table 3.3. Description of the gene hybrids found in *Glomus sp.* 229456 mtDNA.

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

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

These gene hybrid structures are similar to that of the *atp6* gene previously described in the *Allomyces macrogynus* (Figure 3.4, grayscale box), a species that belongs to the basal fungal phylum Blastocladiomycota (Paquin et al. 1994). The same scenario has also been observed in

the *Rhizopus oryzae atp9* and *Mortierella verticillata cox2* genes (Seif et al. 2005). These hybrids contain a carboxy-terminal duplication as well as a mORF encoding an endonuclease, which has been biochemically demonstrated to be responsible for the element mobility. In *Allomyces macrogynus*, the inserted C-terminal was shown to have been recently acquired by HGT based on the divergence in sequence it had with the native C*-terminal, while the latter had a perfect sequence identity with the corresponding gene portion of the closely related species *Allomyces arbusculus*.

2.5.5 Evidences of horizontal gene transfer between *Glomus spp.*

The *Glomus sp. atp6, atp9*, *cox2* and *nad3* native C*-terminals showed higher nucleotide sequence identity to those of *G. irregulare* 494 (91, 98, 93 and 98%, respectively) than their duplicated C-terminal counterparts (64, 71 and 81 and 73%, respectively) (Figures S3.1, S3.2, S3.3, and S3.4 and Tables S3.1, S3.2, S3.3, and S3.4). However at the protein level, the comparison of the C*-terminal amino acid sequences of the *atp6*, *atp9, cox2* and *nad3* genes with the corresponding portion in *G. irregulare* 494 was 100% for *atp6, 94% for atp9, and* 100% for *cox2* and *nad3*. The high sequence identity of the native *Glomus sp.* C*-terminals with *G. irregulare 494,* is in stark contrast to the low similarity observed with the inserted C-terminal portions and points to a recent HGT event, as was described in *Allomyces spp.* (Paquin et al. 1994). However, the HGT hypothesis could likely apply to the *atp6* and *cox2* genes, since their native C*-terminal portion is no longer translated and could undergo rapid divergence. For the *atp9* and *nad3* hybrids, even though it is less parsimonious, the observed sequence divergence between the duplicated portions could have been caused by independent evolution following the mobile element insertion, since both are expressed in the mRNA transcript. It would also be

interesting to see if some of the reported gene hybrids can still accomplish their functions at the protein level, given that the mORF and both C*-terminals are expressed in some cases. They are apparently expressed pseudogenes but post-translational modification mechanisms may be in place to ensure that the resulting protein is functional. We did not find a mORF-less copy of those genes that could have been transferred to the *Glomus sp.* nuclear genome that could explain a pseudogenization in *Glomus sp.* mtDNA. In regards to the HGT hypothesis, and in order to evaluate whether there is a plausible donor for the duplication, we compared the carboxy-terminal sequence of these genes with those in 11 *Glomus spp.* (to avoid redundancy we didn't add the *G. margarita* sequences since they are identical to *G. rosea*) and three phylogenetically related fungal representatives (Figure 3.5). In all four *Glomus sp.* gene hybrids (*atp6, atp9*, *cox2* and *nad3*), the native C*-terminal sequences cluster within the *Glomus spp.* group as expected given the reference phylogeny (Figure 3.5, grayscale box), thereby supporting a recent insertion of the foreign element. The *atp6* gene carboxy-terminal comparison (Figure 3.5A) shows that the mORF-derived C-terminal is related to a *Glomus sp.* isolate DAOM213198 with a moderate 60% bootstrap value. Surprisingly, in *atp9* (Figure 3.5B) the inserted C-terminal is even more distantly related to *Glomus spp.* than to *G. rosea*. In *cox2* (Figure 3.5C) the *Glomus sp.* inserted C-terminal and the more divergent AMF species *G. cerebriforme* are in the same cluster. Finally, the *nad3* C-terminal clustered with *Glomus sp.* 213198, as it was the case for *atp6,* with a 79% bootstrap value (Figure 3.5D). Also, the *nad3* gene shows high variability in length in *Glomus spp.,* due to the insertion of those elements.

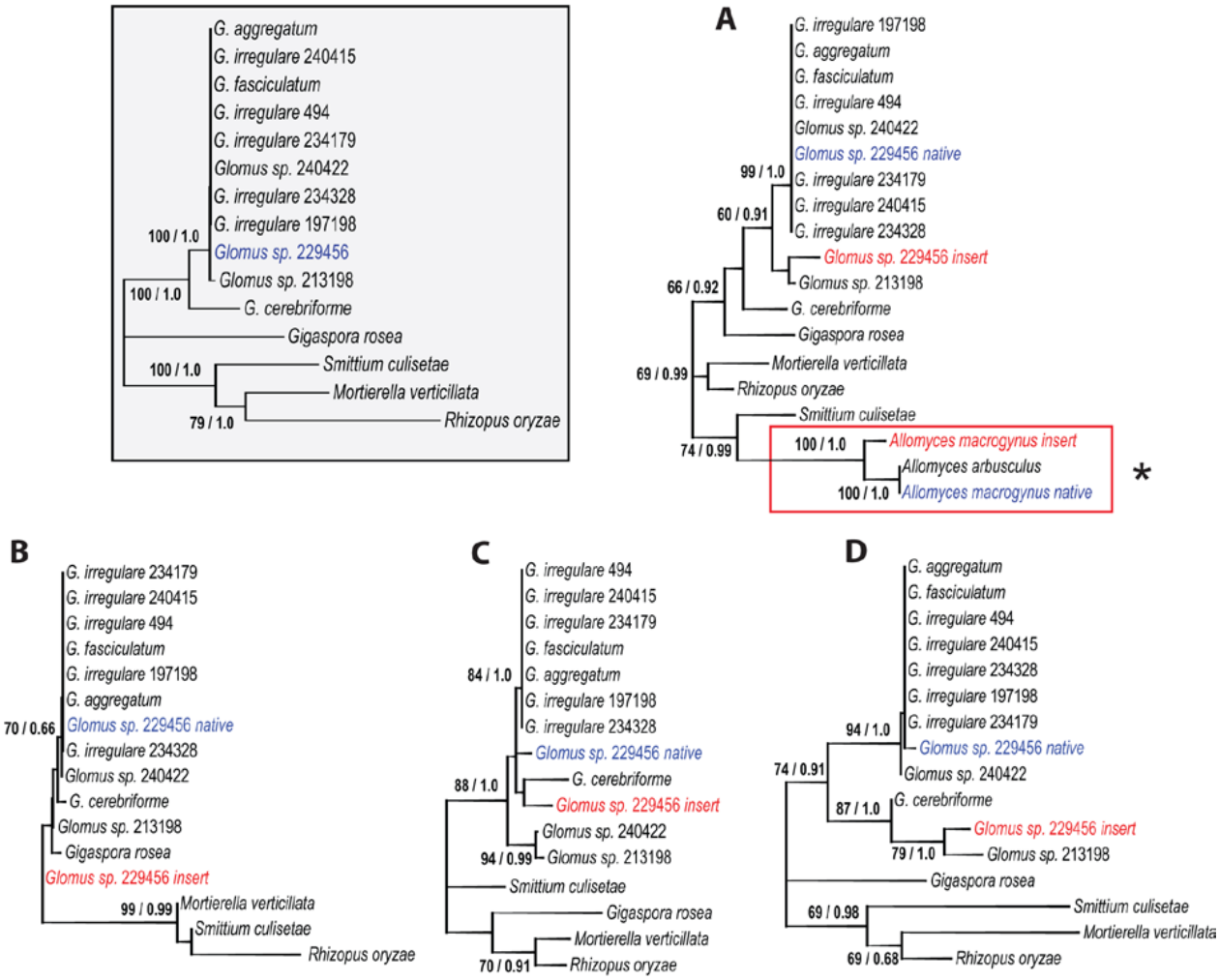


Figure 3.5. Native and inserted C-terminals unrooted maximum likelihood phylogenetic trees. The first number at branches indicates ML bootstrap values with 1000 bootstrap replicates and the second number indicates posterior probability values of a MrBayes analysis with four independant chains. Bayesian inference predicts similar trees (not shown). The concatenated tree of the *atp6*, *atp9*, *cox2* and*nad3* 'core' genes (without the duplicated C*- terminal portion) (1489 alignment positions) of selected AMF representatives (grayscale boxed) are compared with those of the *atp6* (298 alignment positions), where the red box with the asterisk point out to the reference *Allomyces spp.* HGT event (Figure 3.4) (A), *atp9* (51 alignment positions) (B), *cox2* (106 alignment positions) (C) and *nad3* (120 alignment positions) (D) C*-terminals. The *Glomus sp.* native C*-terminals are in blue, while the inserted C-terminals are in red.

In all four cases, the native *Glomus sp.* C*-terminal is nested within the *Glomus spp*. group

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

3.5 Conclusion

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

contributed to AMF evolution is a topic that merits exploration in future studies.

3.6 Acknowledgments

We thank Dr. B.F. Lang for bioinformatics assistance and access to an automated organelle genome annotation software and to Dr. David Morse and Dr. Terrence Bell for comments on the manuscript. We also like to thank Biopterre centre du développement des bioproduits and CRBM for their help.

Presentation of article 3

Numerous mitochondrial genomes of AMF has been sequenced, mapped and deposited in public databases during the past few years. Although genome structure (typically circular) and gene content were similar in all the sequenced mt genomes so far, genome size and genome arrangement were divergent among them. Variance in size of intronic and intergenic regions are the major causes of polymorphism in mt genomes commonly mediated by indels of mobile elements such as open reading frames (*orfs*), plasmid-related DNA polymerase sequences (*dpo*) and short inverted repeats (SIRs) (Formey et al. 2012, Beaudet et al. 2013). Opposing to the available AMF mtDNA, we sequenced mitogenome of an isolate of *Rhizophagus* sp., a close relatives of *Rhizophagus irregularis* reveals the presence of fragmented mtDNA.

In this study, for the first time we report a novel mtDNA organization in the fungal kingdom. Using bioinformatic analyses and experimental evidence, we characterize the unusual multi-chromosomal architecture of mtDNA in *Glomeormycota* (*Rhizophagus* sp. isolate DAOM 213198, a close relative of the model AMF *Rhizophagus irregularis*). This peculiar mtDNA architecture provides the opportunity to assess the evolutionary process on mtDNA in *Glomeromycota* and the mechanism by which the fragmentation of mtDNA occurs. The mt genome of *Rhizophagus* sp. DAOM 213198 provides insights into the role of *dpo* and SIRs in mtDNA structure and evolution.

The Article 3 is accepted for publication in Genome Biology and Evolution. I have taken the leadership in this study by discovering the organization of this mtDNA using bioinformatics followed by PCR and Sangers sequencing. I have done the major part of the sequence analyses, and lab work. F.O. Stefani contributed in qPCR by designing primers, probes and performing the quantifications. I have drafted the manuscript and the final version was completed by M. Hijri.

Chapter 4 - The mitochondrial genome of the glomeromycete *Rhizophagus* sp. DAOM 213198 reveals an unusual organization consisting of two circular chromosomes

Maryam Nadimi, Franck O.P. Stefani, and Mohamed Hijri [*1]

[1] Université de Montréal, Département de sciences biologiques, Institut de recherche en biologie végétale (IRBV), 4101 rue Sherbrooke Est, Montréal, QC, H1X 2B2, Canada.

Published in Genome Biology and Evolution, (2015) 7 (1): 96-105. doi: 10.1093/gbe/evu268

4.1 Abstract

Mitochondrial (mt) genomes are intensively studied in Ascomycota and Basidiomycota, but they are poorly documented in basal fungal lineages. In this study, we sequenced the complete mtDNA of *Rhizophagus* sp. DAOM 213198, a close relative to *Rhizophagus irregularis*, a widespread, ecologically and economical relevant species belonging to Glomeromycota. Unlike all other known taxonomically close relatives harboring a full-length circular chromosome, mtDNA of *Rhizophagus* sp. reveals an unusual organization with two circular chromosomes of 61,964 and 29,078 bp. The large chromosome contained nine protein-coding genes (*atp9*, *nad5*, *cob*, *nad4*, *nad1*, *nad4L*, *cox1*, *cox2*, and *atp8*), small subunit rRNA gene (*rns*), and harbored 20 tRNA-coding genes and 10 *orfs*, while the small chromosome contained five protein-coding genes (*atp6*, *nad2*, *nad3*, *nad6*, and *cox3*), large subunit rRNA gene (*rnl*) in addition to 5 tRNA-coding genes, and 8 plasmid-related DNA polymerases (*dpo*). Although structural variation of plant mt genomes is well documented, this study is the first report of the presence of two circular mt genomes in arbuscular mycorrhizal fungi. Interestingly, the presence of *dpo* at the breakage point in intergenes *cox1-cox2* and *rnl-atp6* for large and small mtDNAs, respectively, could be responsible for the conversion of *Rhizophagus* sp. mtDNA into two chromosomes. Using quantitative real-time polymerase chain reaction, we found that both mtDNAs have an equal abundance. This study reports a novel mtDNA organization in Glomeromycota and highlights the importance of studying early divergent fungal lineages to describe novel evolutionary pathways in the fungal kingdom.

Key words: mitochondrial genome, genome sequencing, basal fungal lineages, fungi, plasmid-like DNA polymerase genes (*dpo*), arbuscular mycorrhizal fungi, Glomeromycota, *Rhizophagus*.

4.3 Introduction

Mitochondria are membrane-bound organelles that are involved in several cell processes such as adenosine triphosphate (ATP) production via oxidative phosphorylation, respiration, RNA maturation, and protein synthesis. Mitochondria are also involved in cell division, growth, and death. They harbor their own genetic material that has evolved from an ancestral prokaryote genome. The endosymbiotic theory (Margulis 1971) suggests that the origin of nuclear genome of eukaryotic cells evolved in parallel to the origin of mitochondrial (mt) genome (Gray et al. 1999; Lang et al. 1999). Structure, size, and even function of mt genomes are variable among eukaryotes (Fukuhara et al. 1993; Drissi et al. 1994; Nosek and Tomaska 2003; Alverson et al. 2011; Nadimi et al. 2012; Beaudet et al. 2013). Previous studies have shown that mitochondria exhibit a large diversity of genome architectures. For example, linear, circular, and fragmented mtDNAs have been described in *Cucumis*, *Phythium*, *Ichthyosporean protists*, *Globodera pallida*, *Pediculus humanus capitis*, *Candida labiduridarum*, *Candida frijolesensis*, and *Brachionus plicatilis* (Martin 1995; Armstrong et al. 2000; Burger et al. 2003; Suga et al. 2008; Alverson et al. 2011; Valach et al. 2011; Shao et al. 2012). Genome reshuffling and evolution of mtDNA structures have been observed in phylogenetically distant eukaryotic lineages as well as in closely related species (Fukuhara et al. 1993; Wilson and Williamson 1997). Conversion of circular genomes to monomeric linear genomes has been shown to occur by an insertion of linear plasmids with inverted terminal repeats (Schnare et al. 1986; Heinonen et al. 1987), resulting in the extension of mt genome size. Another feature of mtDNA is the size variability among eukaryotic lineages, spanning from ~6 kb in *Plasmodium falciparum* (Apicomplexa) to 11.3 Mb in the angiosperm genus *Silene* (Conway et al. 2000; Sloan et al. 2012). Mitochondria without any genes and organisms lacking mitochondria have been reported (reviewed in Keeling and

Slamovits 2004). The gene content of mtDNA also varies broadly from 5 genes in *Plasmodium* (Conway et al. 2000) to 100 genes in jakobid flagellates (Burger et al. 2013), while 40–50 genes are commonly observed in mtDNA of eukaryotes. Gene content, size of introns, intergenic regions, and mobile elements such as open reading frames (*orfs*), plasmid-related DNA polymerase sequences (*dpo*), and short inverted repeats (SIRs) are the major causes of polymorphism in mt genomes of eukaryotes.

Arbuscular mycorrhizal fungi (AMF) are members of the phylum Glomeromycota (Schüssler et al. 2001) and they represent an early-diverging fungal lineage dating back to the Early Devonian (Remy et al. 1994; Redecker et al. 2000). AMF are plant root-inhabiting fungi, where they form mutualistic symbiotic associations with ~80% of vascular plants (Smith and Read 2008). They promote plant growth by enhancing mineral uptake, in particular phosphorus, and protect plants against pathogens by controlling the growth of some soil fungal pathogens or by inducing plant defense responses (Ismail et al. 2011, 2013; Ismail and Hijri 2012). Recently, nuclear and mitochondrial genomics of AMF have been intensively studied (Tisserant et al. 2012, 2013; Halary et al. 2013). The first published mt genome of AMF was *Rhizophagus irregularis* isolate 494, (previously named *Glomus intraradices* and then *Glomus irregulare*) followed by the publication of the mt genomes of 11 taxa belonging to the genera *Rhizophagus*, *Glomus*, and *Gigaspora* (Lee and Young 2009; Formey et al. 2012; Nadimi et al. 2012; Pelin et al. 2012; Beaudet et al. 2013; Beaudet et al. 2013; de la Providencia et al. 2013). AMF identification using the traditional ribosomal DNA markers of the nuclear genomes is uncertain due to high levels of intraspecific variations (Stockinger et al. 2009; Kruger et al. 2012; Schoch and Seifert 2012). Therefore, the publication of mt genomes provides useful data to identify AMF strains. For instance, sequences from intergenic and intronic regions are very divergent, which allows discrimination of closely related isolates (Formey et al. 2012; Beaudet et al. 2013; de la

Providencia et al. 2013).

Mitochondrial genome sequencing provides insights into the mtDNA evolution within Glomeromycota. Indeed, mtDNA structure in Glomeromycota has been shown to undergo different evolutionary mechanisms such as fragmented genes (Nadimi et al. 2012), lateral gene transfer, insertion/excision of mobile elements (Beaudet et al. 2013; Beaudet et al. 2013), and transmission of SIRs (Formey et al. 2012; Beaudet et al. 2013). AMF mt genomes have been invaded by different types of selfish mobile genetic elements (MGEs) or mobilomes, such as homing endonuclease, plasmid-related DNA polymerase (*dpo*), and SIRs (Formey et al. 2012; Beaudet et al. 2013). However, their movement and recombination mechanisms are not clearly understood. MGEs are typically known as DNA fragments encoding enzymes and other proteins that mediate the movement of the related chromosomal segment within genomes (intracellular mobility) or between different individuals (intercellular mobility). SIRs are thought to cause a double-strand breakage (DSB) in close vicinity resulting in initiation of their replication (Rattray et al. 2001). Pairing of two similar SIRs can lead to intrachromosomal homologous recombination resulting in genome rearrangements (Smith 1988; Romero and Palacios 1997).

The plasmid-related DNA polymerase (*dpo*) genes typically found in mitochondria are believed to be of a bacterial origin (Griffiths 1995; Cermakian et al. 1997). Mitochondrial plasmids are small extragenomic mtDNA molecules that can be transmitted vertically and horizontally, increasing the probability of gene transfer between genetically distinct mitochondrial genomes (Yang and Griths 1993). *dpo* genes have been found in mtDNA of many fungi such as Glomeromycota. Using comparative mt genomics, Beaudet et al. (2013) reported evidence of *dpo*-mediated interhaplotype recombination leading to mt genome rearrangement in *Rhizophagus* sp. DAOM 240422. This study revealed the presence of highly similar plasmids in distantly related fungal lineages, supporting horizontal gene transmission of these elements.

However, it is not yet clear how *dpo* sequences are integrated into mtDNA and how they move within mtDNAs.

In this study, we report a novel mtDNA organization in the *Rhizophagus* sp. isolate DAOM 213198, a close relative of the model AMF *R. irregularis*. Using bioinformatic analyses and experimental evidence, we characterize the unusual multichromosomal architecture of mtDNA in Glomeormycota. This peculiar mtDNA architecture provides the opportunity to assess the evolutionary process on mtDNA in Glomeromycota and the mechanism by which the fragmentation of mtDNA occurs. The mt genome of *Rhizophagus* sp. DAOM 213198 provides insights into the role of *dpo* and SIRs in mtDNA structure and evolution.

4.4 Materials and methods

4.4.1 Fungal material and DNA extraction

Rhizophagus sp. DAOM 213198 (synonym MUCL 43203) was obtained from the National Mycological Herbarium (DAOM), Ottawa, ON, Canada. This isolate was isolated from the perennial grass Agrostis gigantean collected from Quetico Provincial Park, ON, Canada, in 1989. Spore morphology showed some similarities with *R. irregularis* (supplementary fig. S4.1, Supplementary Material online). The isolate DAOM 213198 is closely related to *R. irregularis* (molecular virtual taxon VTX00114, accession number: AM849267) because the pairwise alignment of their 18S rDNA sequences has 99% of similarity (13 different nucleotide substitutions out of 1,716 bp).

The isolate DAOM 213198 was cultivated *in vitro* on a minimal (M) medium with carrot roots transformed with Agrobacterium rhizogenes. The growing medium containing mycelium and spores was dissolved in extraction buffer (0.82 mM sodium citrate and 0.18 mM citric acid).

DNA was extracted from spores and hyphae using the DNeasy Plant Mini Kit (Qiagen, Toronto, ON) according to the manufacturer's recommendations, except that spores and hyphae were crushed using a micropestle in a 1.5-ml tube containing 400 µl of AP1 buffer.

4.4.2 Sequencing and Bioinformatic Analyses

Total DNA from *Rhizophagus* sp. DAOM 213198 was sequenced using Roche 454 GS FLX Titanium. The full sequence run and subsequent Sanger sequencing were performed at the Genome Quebec Innovation Center (McGill University, Montreal, QC). Reads were de novo assembled with Newbler software (Roche, Version 2.9) at the Genome Quebec Innovation Center. Gene annotation was performed automatically with MFannot (http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl, last accessed December 16, 2014), followed by manual inspection and editing as described in Nadimi et al. (2012). The mtDNA of *R. irregularis* DAOM 234179 (KC164354) was used as a reference for comparative analyses. This isolate has been chosen as a reference because of its common gene synteny among Glomeromycetes (Beaudet et al. 2013). Polymerase Chain reaction (PCR) and Sanger sequencing were performed to fill the gaps and join mtDNA contigs. The annotated sequences of the mt circular chromosomes of *Rhizophagus* sp. DAOM 213198 were deposited in GenBank under accession numbers KF591215 and KF591216, respectively. mtDNA circular maps were built using OGDRAW Version 1.2 software (Lohse et al. 2013). Sequence and comparative analyses were performed using the National Center for Biotechnology Information (NCBI) genomic database and multiple sequence alignment. Sequences were compared at the amino acid level due to high mutation and substitution rates in intergenic regions and dpo-like sequences. A search with TBLASTX was used to find the translated nucleotide database in NCBI using a minimum E-

value cutoff of 1×10^{-10} and 50% minimum identity.

4.4.3. Polymerase Chain Reactions

Conventional and long-range PCRs were used to sequence gaps between contigs, to validate gene synteny, to demonstrate the circularity of the small mt chromosome (primers are listed in table 4.1), and to correct potential pyrosequencing errors in homopolymer stretches. Conventional PCRs were performed in a final volume of 50 µl containing 1X PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM of each deoxynucleotide triphosphate, 0.5 µM of each primer, and 1 unit of Platinum Taq DNA Polymerase (Life Technologies, Burlington, ON). Cycling parameters were as follows: 94 °C for 90 s, followed by 38 cycles of 94 °C for 1 min, annealing temperature (table 4.1) for 30 s, 72 °C for 90 s, and a final elongation at 72 °C for 5 min. Long-range PCR reactions were also performed to attempt to assemble the last two mt contigs. Long-range PCR reactions were done in a final volume of 50 µl using TaKaRa LA PCR Amplification Kit (Clontech, Mountain View, CA) according to the manufacturer's instructions or using the PCR protocol for LongAmp Hot Start Taq DNA Polymerase (New England Biolabs). PCRs were run in an Eppendorf Mastercycler ProS (Eppendorf, Mississauga, ON). PCR products were separated by electrophoresis in a 1% or 0.8% (w/v) agarose gel, stained with GelRed Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA) and visualized with a Gel Doc System (BioRad, Mississauga, ON).

Table 4.1. Primers and probes designed for the real-time qPCR assays and long range PCR primers used to validate the circularity of the two mtDNAs in *Rhizophagus* sp. DAOM 213198. Primer direction (F, forward and R, reverse), sequence, PCR product size in base pair (bp) and melting temperature in degree Celsius (Tm) are indicated.

Region and primer name	Primer direction	Sequence (5'-3')	Size (bp)	Tm (°C)
cox1	F	CTTCGCAGGTTGGAATATG	5,690	60
cox2	R	CATCCAGCCAACACCTAGTAG		59
rnl	F	ATAGGACCTGGGTGTAATAGC	3,309	58
atp6	R	GCCTATATTAGTCAGCCCA		56
8.1	F	GGACAAATACGCTACTCTTATAGATTCAGAAC	9,911	65.1
6.2	R	GCTTCAAACATTCTGATGAGGCTAAAGA		65.6
7.1	F	GGAGAGTAGGTTTTAAACAGTGTCTAATTTCTG	9,100	65.6
8.2	R	GTTCTGAATCTATAAGAGTAGCGTATTTGTCC		65.1
6.1	F	TCTTTAGCCTCATCAGAATGTTTGAAGC	10,160	65.6
7.2	R	CAGAAATTAGACACTGTTTAAAACCTACTCTCC		65.6
	F	TCTTCGTGCTTCTGACTAC		55.5
nad4	R	GTCTCTATCACAAGTAGCGA	125	54.4
	Probe	TCCGGTAATATTTCCGCTGTCAA		65
	F	ATGAAGTTTCAACCTCATCCTTATC		62.2
cox3	R	TCCTCCATGACCATATCCATG	126	64.5
	Probe	AGTAGAACCTTCACCATGGCCTCTAGCA		71.7
	F	GACTTTACGTGGTTCTAGTTGTTAG		60.6
rnl intron 4	R	CTACCCTAGTAAGTAAGGGTTTGG	150	60.9
	Probe	TGAAACAATTGGGTTCAAATCAAGGGTTGT		67.1

4.4.4 Quantitative real-time PCR

The abundance of each mt genome in DAOM 213198 was assessed by quantitative real-time PCR (qPCR). Four specific TaqMan assays (table 4.1) targeting intron 4 of *rnl* and *cox3* genes for the small mt chromosome and intron 7 of *cox1* and *nad4* genes for the large mt chromosome were designed to quantify the abundance of each chromosome. TaqMan probes labeled with 5′

FAM fluorophore and 3′ BHQ-1 quencher (Alpha DNA, Montreal, QC) and primers were designed using the plugin Primer3 (Untergrasser et al. 2012) implemented in Geneious Version 7.1.4 (BioMatters, Auckland, NZ). The cycling parameters for qPCR were 95 °C for 3 min, followed by 40 cycles of 95 °C for 15 s and the final elongation at 60 °C for 1 min. qPCRs were performed in a final volume of 20 µl containing 10 µl of iTaq Universal Probes Supermix (Bio-Rad Laboratories, Mississauga, ON), 0.65 µM of each primer, 0.08 µM of TaqMan probe, and 2.7 µl of DNA template. Four and three 10-fold serial dilutions were used as DNA templates for DAOM 213198 and 197198, respectively. Isolate DAOM 197198 was used for comparison purposes. Absolute quantification was performed using six 10-fold serial dilutions of circular plasmids. DNA was quantified using the Qubit dsDNA BR and HS Assay Kits (Life Technologies) and the Qubit 2.0 Fluorometer (Life Technologies) according to the manufacturer's instructions. Each sample was amplified in triplicate in a ViiA 7 Real-Time PCR System (Life Technologies). PCR efficiency was calculated by converting the slope produced by the linear regression of the curves to percentage efficiency using the formula: Efficiency = −1 + 10 (−1/slope).

4.5 Results and Discussion

4.5.1 Description of *Rhizophagus* sp. mtDNA

A total of 223,988 reads were produced with an average size of 400 bp. *De novo* DNA sequence assembly of *Rhizophagus* sp. produced 8 contigs ranging from 4,014 to 33,711 bp, containing all the expected mtDNA genes previously recorded in isolates closely related to R. irregularis. The sequence coverage of mtDNA contigs ranged between 10 × and 40 ×. The mtDNA sequence of *R. irregularis* DAOM 234179 for which SIRs and plasmid-related DNA

polymerase (*dpo*) sequences were annotated, was used as a reference for gene synteny and for designing primer sets at the termini of each contig to complete contig assemblies by Sanger sequencing. The eight contigs were assembled in two contigs of 29 and 56 kb. Many attempts were made to assemble the 29-kb contig flanked by atp6 and rnl and the 56-kb contig flanked by *cox2* and *cox1* using the primers listed in table 4.1.

Using conventional and long PCRs, no amplification was observed for any of the four possible primer combinations while positive controls produced fragments of the expected size (data not shown). We therefore attempted to join the termini of each contig with the primers designed at their termini (fig. 4.1A). The combination of the forward rnl primer with the reverse atp6 primer produced an amplicon of 3,309 bp, while a PCR product of 5,690 bp was obtained using the forward *cox1* primer with the reverse *cox2* primer (fig. 4.1 and table 4.1). Sanger sequencing of these PCR products confirmed the circularity of the ~29 and ~56 kb contigs. To double-check mtDNA circularity, the small mt chromosome was fully amplified using overlapping long-range PCR. Three positions of the small mt chromosome were selected to design three forward primers and their complementary sequences were used as reverse primers (table 4.1 and fig. 4.2B). The three primer combinations successfully amplified three fragments of the expected sizes covering the entire small chromosome (fig. 4.2C). Taken together, these results confirmed that the mt genome of DAOM 213198 is made of two circular chromosomes (fig. 4.2), an unusual structure because previously sequenced mt genomes within AMF exhibited a single circular chromosome.

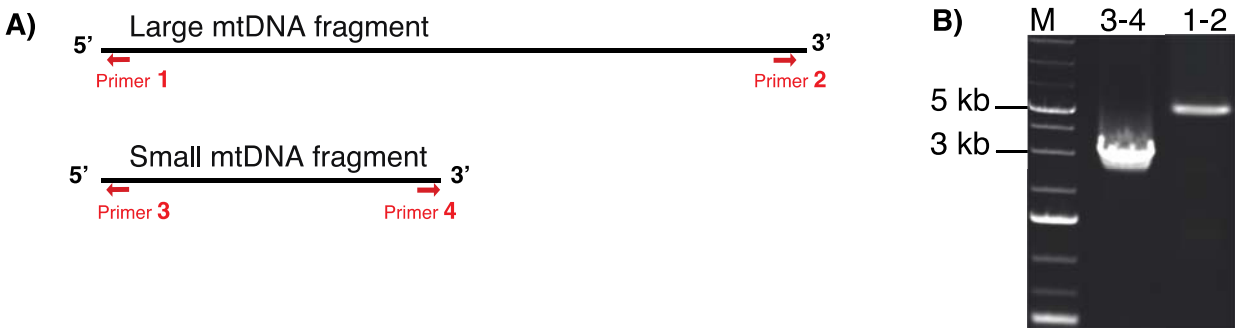


Figure 4.1. A, Schematic representation of the two mtDNA contigs obtained after 454 reads assembly of *Rhizophagus* sp. DAOM 213198.

Red arrows represent the position and orientation of primers shown in Table 1. B, Agarose gel electrophoresis showing long range PCR amplification patterns using primers designed in the termini of each contig. The sizes of the amplicons were 5,690 and 3,309 for '1-2' and '3-4' primer combinations, respectively. M, molecular marker.

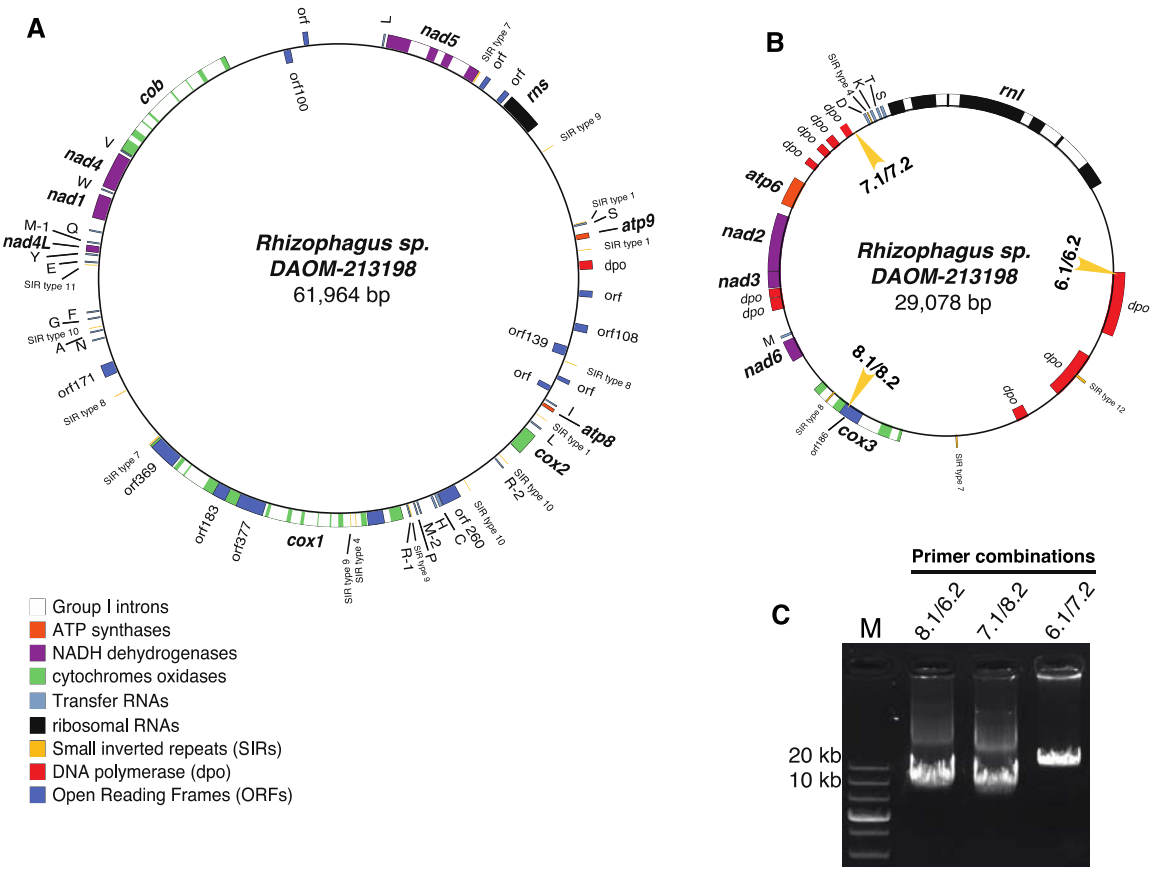


Figure 4.2. mtDNA circular maps of *Rhizophagus* sp. DAOM 213198.

The mt genome consists of two circular-mapping chromosomes (A, large mtDNA chromosome; B, small mtDNA chromosome). The outer and inner boxes show genes that are transcribed in a clockwise and counter clockwise direction, respectively. Gene color codes are indicated. C, PCR amplifications of the entire small mtDNA using primers indicated by yellow arrows in panel B. Three PCR bands of the expected sizes of 9,911 bp, 9,100 bp and 10,160 bp were recovered respectively using primer combinations 8.1/6.2, 7.1/8.2 and 6.1/7.2 (Table 1). M, molecular marker.

Prior to completing the whole mt maps, the presence of ~6 kb fragments between atp8 and atp9 was validated by PCR and Sanger sequencing approaches, finalizing the larger chromosome size at 61,942 bp. Compared with closely related species, this region contained unexpected dpo-like fragments and previously unidentified sequences based on querying GenBank database.

These fragments are the major cause of mt genome expansion in the size of *Rhizophagus* sp. DAOM 213198 (figs. 4.1 and 4.2). Together, the two mtDNA chromosomes represented 91,042 bp, which is the largest mt genome sequenced so far in Glomerales order, while *Gigaspora rosea* has the largest mt genome recorded in Glomeromycota (97.3 kb; Nadimi et al. 2012). Together, the two mtDNAs of *Rhizophagus* sp. DAOM 213198 contained the 41 mt genes recorded in *R. irregularis* DAOM 234179, consisting of two ribosomal ribonucleic acids (rRNAs), 14 protein-coding genes, and 25 transcribed ribosomal ribonucleic acid (tRNA)-coding genes. The 29-kb mtDNA chromosome contained *rnl*, *atp6*, *nad2*, *nad3*, *nad6*, and *cox3* in addition to five tRNAs and eight plasmid-related DNA polymerases (*dpo*). The 62-kb chromosome contained *rns*, *atp9*, *nad5*, *cob*, *nad4*, *nad1*, *nad4L*, *cox1*, cox2*,* and *atp8* and harbored 20 tRNAs and 3 *dpo* sequences. Both mt chromosomes showed a low Guanine-Cytosine content of 36.5%.

4.5.2 Mitochondrial Genome Conversion

Although the gene synteny in *Rhizophagus* sp. DAOM 213198 is similar to *R. irregularis* DAOM 234179 in nonrecombined regions, putative division of the ancestral circular mtDNA created two novel intergenic regions between *rnl-atp6* and *cox1-cox2* (figs. 4.3 and 4.4). The comparative analysis of AMF mt genomes suggests the presence of hot-spot regions subjected to recombination mediated by the integration of dpo sequences, such as intergenic regions of *cox3-rnl*, *cox1-nad4L*, and *rns-atp9* (Beaudet et al. 2013).

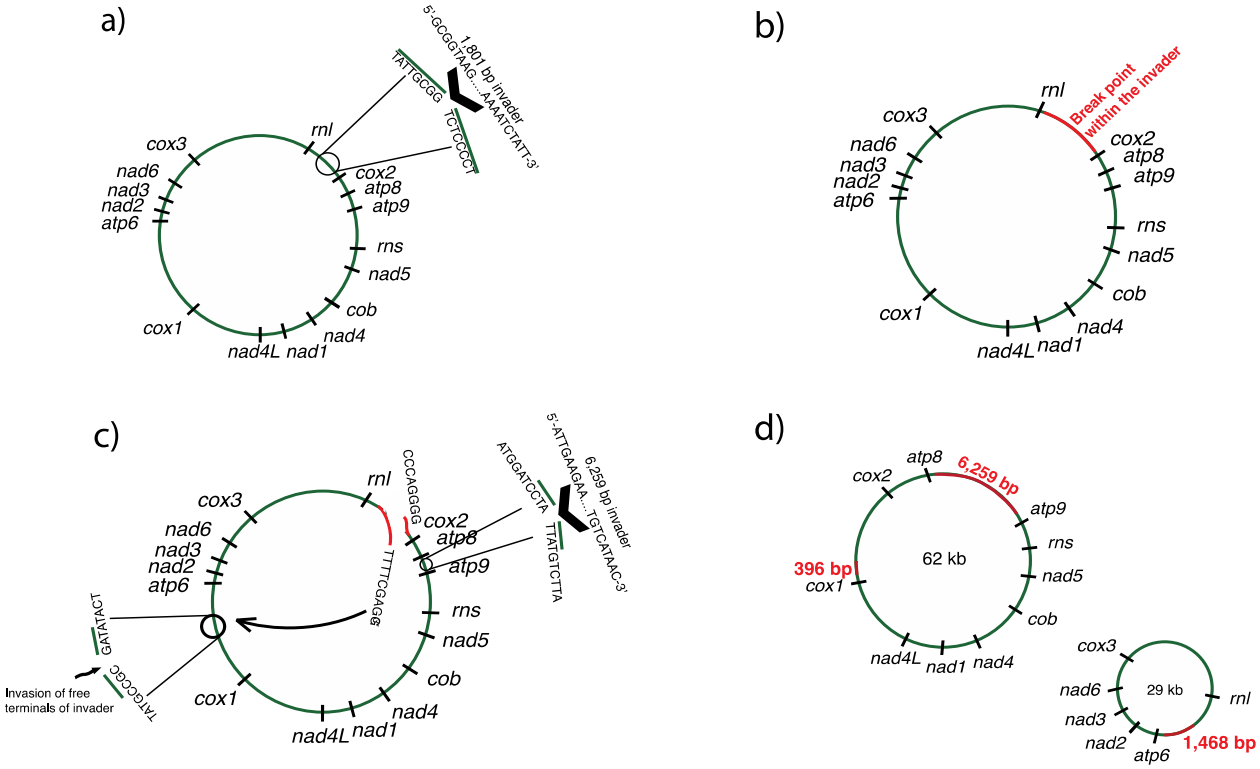


Figure 4.3. Hypothetical pathway of mtDNA conversion in Rhizophagus sp. DAOM 213198. a) Insertion of 1,801 bp *dpo*-like sequence in intergene of *rnl-cox2* (or possibly in *atp6-cox1*) in ancestral mtDNA. The circle projection shows the break points with their corresponding terminal sequences while the black arrow-head shows the insertion point. b) Putative ancestral mtDNA shows insertion of an invader sequence (red). c) Breakage of inserted fragment (red) and ligation of its 3' end to N-terminal sequence of an inserted sequence intergene *cox1-atp6* followed by subsequent ligation of 3' end of *cox1-atp6* inserted sequence with N-terminal of the *rnl-cox2* broken sequence. Another insertion of a 6,259 bp sequence occurred in *atp8-atp9* intergene. d) Division of the ancestral mtDNA into two circular-mapping mtDNA in *Rhizophagus* sp. DAOM 213198. Red fragments represent the inserted sequences of 396 bp and 6,259 pb for the 62 kb fragment and 1,468 bp for the 29 kb fragment.

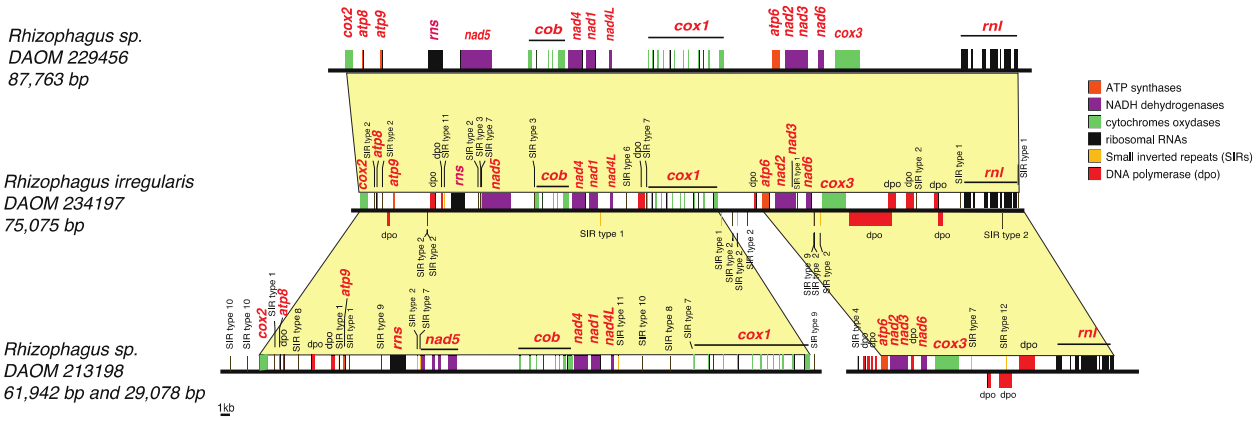


Figure 4.4. Linear representation of mtDNAs of *R. irregularis* DAOM 234179, DAOM 229456 and *Rhizophagus* sp. DAOM 213198 showing their genome synteny.
Yellow projections represent mtDNA regions with the same gene order.

The comparative sequence analyses of reshuffled intergenic regions of *atp9-atp8*, *rnl-atp6*, and *cox1-cox2* between the mitochondrial deoxyribonucleic acids of DAOM 213198 and the close relative *R. irregularis* DAOM 234179 suggest that conversion and recombination could act as mechanisms by which the fragmentation of mtDNA occurred in DAOM 213198 (fig. 4.3). Interestingly, an intergenic region between *cox1* and *cox2* has never been found in any other Glomeromycota sequenced so far. The nucleotide sequences of both termini of this unique intergene matched with the putative ancestral intergenes of *rnl-cox2* and *cox1-atp6* in *R. irregularis* DAOM 234179. This supports the idea that recombination could be initiated by a DSB followed by translocation and reshuffling, resulting in fragmentation of the ancestral mtDNA into two mtDNA circular molecules (fig. 4.3). Comparison between *rnl-atp6* and *cox1-cox2* intergenes in DAOM 213198 and *rnl-cox2* and *cox1-atp6* intergenes in *R. irregularis* DAOM 234179 suggests a *dpo*-like sequence insertion into their common ancestral mtDNA. The inserted *dpo*-like sequence contains some conserved *dpo* domains (which are also found in other mt genomes within Glomeromycetes) flanked by nonidentified sequences according to a

TBLASTX search (fig. 4.3 and supplementary fig. S4.2, Supplementary Mateiral online). Conserved domains of *dpo*-like sequences at the amino acid level are shown in supplementary figure S4.3, Supplementary Mateiral online. Figure 4.3 shows the hypothetical mechanism of the mtDNA fragmentation in DAOM 213198: An insertion of a dpo sequence could have occurred in the ancestral intergenic region of *rnl*-*cox2* or *cox1*-*atp6*, followed by a breakdown and disintegration of the inserted *dpo* sequence. Translocation and ligation of breaking sequence termini of *rnl* to *atp6* and *cox1* to *cox2* resulted in the formation of 29- and 62-kb mtDNA circular chromosomes, respectively (fig. 4.3). The presence of *dpo* fragments in this region strongly suggests a *dpo*-mediated recombination as reported previously (Beaudet et al. 2013).

However, the origin of *dpo* domains flanking unknown sequences and their replication and movement mechanisms remain unclear. Comparative genomics analysis between the several AMF mtDNAs in Beaudet et al. (2013) and DAOM 213198 suggests that some intergene regions can be considered as putative hot-spot regions for recombination. Supplementary figure S4.2, Supplementary Material online summarizes putative recombination hot spots by *dpo*-like sequences (containing the conserved domains) that could impact mt genes reshuffling in *Rhizophagus* sp. DAOM 213198 and *R. irregularis* DAOM 234179. This comparative mtDNA analysis provided some clues about the potential mechanisms of genome conversion occurring in the isolate of *Rhizophagus* sp. DAOM 213198.

4.5.3 Short inverted repeats (SIRs) mediate recombination in mtDNA

Recombination in mtDNA can also be mediated by short inverted or palindromic repeats, as previously demonstrated (Rattray et al. 2001, 2005). SIRs recorded in DAOM 213198 and in all mtDNAs of *Rhizophagus* spp. publically available (Formey et al. 2012; Beaudet et al. 2013) were

analyzed to investigate their role in mtDNAs and their dispersal within and among mt genomes in *Rhizophagus*. On the basis of their DNA secondary structure, a total of 13 types of SIRs were recorded in the mtDNAs of DAOM 213198 (fig. 4.5 and supplementary fig. S4.4, Supplementary Material online). The distribution patterns of SIR types, such as types 1 and 2 in *R. irregularis* DAOM 234179, suggest potential repeat-mediated recombination (fig. 4.4 and table 4.2). However, the decreased number of some SIRs' types in mtDNAs of DAOM 213198 that were frequently observed in DAOM 234179 (table 4.2 and fig. 4.5) supports their deletion through repeat-mediated recombination as previously reported (Marechal and Brisson 2010; Davila et al. 2011). It remains unclear how such SIR replications and dispersals take place in mt genomes. Presumably, SIRs tend to invade intergenic regions and introns that are under low selective pressure.

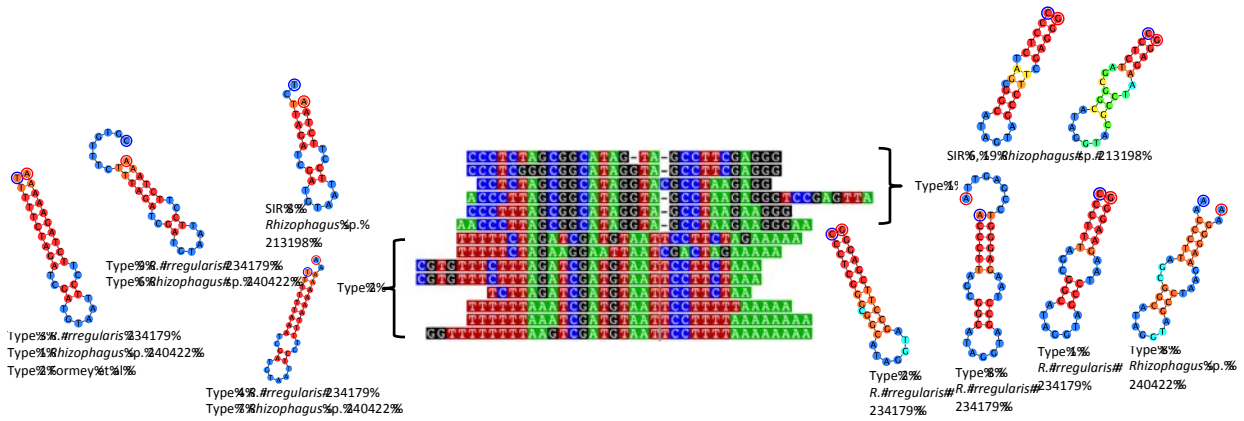


Figure 4.5. Multiple sequence alignment of the most common identified short inverted repeats (SIRs).

Multiple sequence alignment of 13 types of SIRs found in *R. irregularis* DAOM 234179, *Rhizophagus* sp. DAOM 240422, DAOM 213198 and those reported by Formey et al. (2012). SIRs types 1 and 2 are the most frequent SIRs found in all AMF mtDNA analyzed in this study. Secondary structures were predicted based on the energy model of (Mathews, Disney et al. 2004) and (Andronescu, Condon et al. 2007). The minimum free energy (MFE) structure of hairpins is

coloured according to the base-pairing probabilities (red: high; green: mid; blue: low). Blue and red circles around nucleotides represent the beginning and the end of molecules, respectively.

Table 4.2. Distribution of small inverted repeats (SIRs) found in *R. irregularis* DAOM 234179 and *Rhizophagus* sp. DAOM 213198 depending on categorized types and genome localization.

R. irregularis DAOM 234179	
SIR subtypes	Genome localization
T1	*cox1-atp6*
	cox1-atp6
	nad3-nad6
	cox3-rnl
	rnl-cox2
	atp8-atp9
	nad5 intron
	nad1-nad4l
T2	*cox1-atp6*
	cox1-atp6
	nad6-cox3
	nad6-cox3
	cox3-rnl
	cox2-atp8
	atp9-rns
	atp9-rns
	rns-nad5
	cob intron
T3	*rns-nad5*
	nad5-cob
T6	*nad4l-cox3*
T7	*rns-nad5*
	nad4l-cox3
T8	*nad6-cox3*
T11	*atp9-rns*

Rhizophagus sp. DAOM 213198	
SIR subtypes	Genome localization
T1	*atp8-atp9*(2)*
	atp9-rns(2)
	cox2-atp8(2)
T2	*rns-nad5*(2)
T4	*cox1* intron(2)
	rnl-atp6(1)*
T7	*rns-nad5*(2)
	nad4l-cox1(2)
	cox3-rnl(1)
T8	*nad4l-cox1*(2)
	atp8-atp9(2)
	cox3 intron(1)
T9	*atp9-rns*(2)
	cox1 intron(2)
	cox1-cox2(2)
T10	*nad4l-cox1*(2)
	cox1-cox2(2)
	cox1-cox2(2)
T11	*nad4l-cox1*(2)
T12	*cox3-rnl*(1)

4.5.4 Relative quantification of mtDNA chromosomes

Real-time quantitative assays showed that the two mtDNAs had a similar abundance (supplementary table S4.1, Supplementary Material online). This result is surprising because smaller mt genomes usually replicate faster than larger mt genomes (Birky 2001; Schoch and Seifert 2012). A putative mechanism controlling mtDNA replication might occur in the isolate DAOM 213198, resulting in an equal abundance of both mt chromosomes. This has been shown previously (Birky 2001; Schoch and Seifert 2012) where it was found that replication of mtDNAs could be affected by not only the rate of initiation and the rate of completion of mtDNA synthesis but also by the mechanisms that regulate the overall mass of each mtDNA.

4.5.5 Mitochondrial inheritance

Mt inheritance and dynamics have been intensively studied in many eukaryotes such as mammals, plants, and yeasts (Birky 2001). However, these aspects have not been investigated in Glomeromycota, probably because of their obligate biotrophic lifecycle and their slow growth. Yet, such information about the mitochondrial dynamics and inheritance could give us insights into the evolution of mtDNAs of *Rhizophagus* sp. DAOM 213198. Mt division could produce mitochondria containing both mtDNAs or one of each mtDNA as shown in supplementary figure S4.5, Supplementary Material online. Mitochondria and their mtDNAs could be randomly passed to the progeny during cell division, as reported by Birky (2001). However, mitochondria are able to undergo fusion followed by fission, thereby regulating their mtDNA segregation (supplementary fig. S4.5, Supplementary Material online). If such a mechanism of inheritance regulation did not exist, this could lead to forming nonfunctional mitochondria (Chan 2006; Chen and Chan 2009).

Mt genes encoding essential enzymes and subunits for oxidative phosphorylation pathways for energy (ATP) synthase are located in two mtDNAs in *Rhizophagus* sp. DAOM 213198. It is likely that mitochondria lacking one of the two mtDNAs could not be functional. The isolate DAOM 213198 could have a mechanism that controls mtDNA dynamics and inheritance. Without this mechanism, mitochondria with incomplete mtDNA sets could lead to nonfunctional cells (Benard and Karbowski 2009).

4.6 Conclusions

We have documented an unusual mtDNA organization in the isolate *Rhizophagus* sp. DAOM 213198, a close relative to the model species *R. irregularis*. This novel mtDNA feature leads to a new level of understanding of mt genome evolution in eukaryotes. The mtDNA comparative analyses between close relatives of *Rhizophagus* spp. shows that *dpo*-like sequences and SIR-mediated recombinations not only enhance mtDNA reshuffling but could also lead to fragmentation, impacting mobile element dynamics in mt genomes. Mt genomes show a high potential for developing molecular tool kits in order to discriminate isolates and closely related taxa, and to monitor gene exchange and recombination among isolates. Rearrangement of genes and intergenic regions has potential implications in studying population genetics, ecology, and functions of Glomeromycetes in ecosystems. However, the mtDNA organization in two circular chromosomes found in DAOM 213198 raises fundamental questions about their replication and inheritance compared with other AMF harboring one single circular chromosome with a full set of genes. Further investigations are needed to advance our understanding of the evolution of mtDNA in eukaryotes, particularly in basal fungal lineages.

4.7 Acknowledgments

This work was supported by The Natural Sciences and Engineering Research Council of Canada (RDCPJ 395241-09 and discovery grant 328098-2012 to M.H.). Our thanks to Dr B.F. Lang for bioinformatics assistance and access to an automated organelle genome annotation software, to Dr Y. Dalpé for providing DAOM 213198 strain and associated images in supplementary figure S4.1, Supplementary Material online, to Dr D. Beaudet and Dr F. Bou Daher for their comments on the manuscript, and to Dr K. Fisher-Favret for English editing and comments.

Presentation of article 4

The increasing number of sequenced AMF mitochondrial genes and mitogenomes shed light on the importance of this genomic pool for evolutionary analyses of AMF. However, evolutionary signal and phylogenetic power of mt genes have not been evaluated in AMF. Therefore, nobody can truly and statistically claim which mt single gene or subset of genes could be the most proper candidate for phylogenetic and evolutionary analyses. Moreover, it is important to test whether mtDNA could complement other identification methods to identify and characterize AMF at low taxonomic levels (isolate and species level).

In this study, we sequenced and mapped mtDNA of two more AMF species *Rhizophagus irregularis* and *Glomus aggregatum* to implement the mtDNA datasets. The aim was to perform mitogenomics comparative analyses and to assess mt genes phylogenetic signals and potential use to infer phylogenies in AMF. We used two methods of phylogenetic analyses and *in silico* mitogenomic comparative analyses.

The **Article 4** is currently in preparation. I have designed the experiment and performed all bioinformatics analyses as well as phylogenetic and statistical analyses. I have written the complete manuscript which is corrected and organized by M. Hijri.

Chapter 5 - Mitochondrial comparative genomics and phylogenetic signal assessment of mtDNA among arbuscular mycorrhizal fungal taxa

Maryam Nadimi, Laurence Daubois and Mohamed Hijri

Université de Montréal, Département de sciences biologiques, Institut de recherche en biologie végétale (IRBV), 4101 rue Sherbrooke Est, Montréal, QC, H1X 2B2, Canada.

*Corresponding author

Submitted in Molecular Phylogenetics and Evolution.

5.1. Highlights

• Two mitochondrial genomes (mtDNA) of arbuscular mycorrhizal fungi were sequenced and annotated.

• Mitochondrial comparative genomic analysis was performed on 16 Glomeromycotan complete mtDNA.

• Phylogenetic signal of individual, subset and complete set of mt genes was assessed among Glomeromycota taxa.

• mtDNA comparative genomics unraveled the intraspecific diversity of *R. irregularis* species complex.

Keywords: Arbuscular mycorrhizal fungi, Fungi, Mitochondrial genome, Genome evolution, Comparative mitogenomics, Phylogenetic analysis, SH-test, Supergene.

5.2. Abstract

Mitochondrial (mt) genes, such as cytochrome C oxidase genes (*cox*), have been widely used for barcoding in many groups of organisms, although this approach has been less powerful in the fungal kingdom due to the rapid evolution of their mt genomes. The use of mt genes in phylogenetic studies of Dikarya has been met with success, while basal fungal lineages remain less studied, particularly the arbuscular mycorrhizal fungi (AMF). Considering the key roles played by AMF in natural environments, it is critical to develop AMF-targeted molecular toolkits, and to use these to describe the population genetics, community ecology, and phylogeny of AMF. Advances in next-generation sequencing have substantially increased the number of publically available mtDNA sequences for the Glomeromycota. As a result, comparison of

mtDNA across key AMF taxa can now be applied to assess the phylogenetic signal of individual mt coding genes, as well as concatenated subsets of coding genes. Here we show comparative analyses of publically available mt genomes of Glomeromycota, augmented with two mtDNA genomes that were newly sequenced for this study (*Rhizophagus irregularis* DAOM240159 and *Glomus aggregatum* DAOM240163), resulting in 16 complete mtDNA datasets. We assessed the phylogenies inferred from single mt genes and complete sets of coding genes, which are referred to as "supergenes" (16 concatenated coding genes), using Shimodaira-Hasegawa tests, in order to identify genes that best described AMF phylogeny. We found that *rnl*, *nad5*, *cox1,* and *nad2* genes, as well as concatenated subset of these genes, provided phylogenies that were similar to the supergene set. This mitochondrial genomic analysis was also combined with principal coordinate and partitioning analyses, which helped to unravel certain evolutionary relationships in the *Rhizophagus* genus and for *G. aggregatum* within the Glomeromycota.

5.3. Introduction

Elucidating the population structure and inferring the evolutionary histories of species are critical towards understanding their role in ecosystems, and for microorganisms, this generally requires a molecular approach. Developing effective methods, genetic markers, and discriminating criteria are key to phylogenetic analysis and species diagnosis. Mitochondrial (mt) genes and genomes have been used extensively in phylogenetic and population genetic studies of a wide range of organisms. Typical mt genomes comprise a single molecule of DNA, which is usually circular and contains all mt genes, while the number of mt genes varies among species. It has often been assumed that the phylogenetic signal of each mt gene is identical, or highly similar, due to the physical proximity of these genes which lie within the same mtDNA molecule (Avise

1994). However, mtDNA analyses in many organisms have suggested divergence in the phylogenetic signal strength of mt genes among and within species (Duchene et al. 2011). Several studies have indicated that sequencing of complete mtDNA molecules, referred to as "supergenes" (i.e. complete set of mt protein coding genes), provides the highest phylogenetic resolution among studied organisms (Duchene et al. 2011; Fenn et al. 2008; Havird and Santos, 2014). In addition, recent developments in next-generation sequencing technologies have generated a huge number of mtDNA datasets across a broad range of organisms, in particular in the kingdom Animalia (Duchene et al. 2011; Havird and Santos, 2014). These publically available mtDNA datasets offer opportunities to perform large-scale comparative mt genomic analyses, to infer phylogenetic relationships (for identification purposes) and to compare the phylogenetic signals of complete sets of mt genes, or of concatenated subsets of individual mt genes. Numerous phylogenetic studies have been performed on single mt genes such as cytochrome C oxidase subunit 1 (*cox1*) (Borriello et al. 2014; Hebert et al. 2003; Ratnasingham and Hebert 2007; Roe and Sperling 2007) and the mitochondrial large subunit gene (*rnl*) to infer evolutionary relationships between organisms (Borstler 2008; Bruns 1998). However, the phylogenetic signals produced by these extensively studied genes have not often been compared with the complete set of mt genes or with other mt gene subsets. Such comparative analyses are key to accurately assessing the evolutionary relationships of mt genomes within a given group of organisms, and for identifying genes with strong phylogenetic signals that can be used to discriminate between closely related taxa. For instance, the arbuscular mycorrhizal fungi (AMF), a group of closely related symbionts of plant roots, are notoriously difficult to characterize and distinguish through molecular methods. A more comprehensive mitochondrial-based approach could provide a solution to this issue.

The arbuscular mycorrhizal fungi are among the most abundant eukaryotic symbionts of plants, and not only improve plant mineral uptake and enhance plant productivity and diversity (Smith and Read 2008), but also help in the control of plant pathogens (Azcón-Aguilar and Barea 1997; Garcia-Garrido and Ocampo 2002; Ismail and Hijri 2012; Ismail et al. 2011) and stimulate the fitness of plants in polluted environments (Hildebrandt et al. 1999; Zhang et al. 2014). The AMF constitute an early diverging fungal lineage, whose mycelia are typically formed by coenocytic branching hyphae (i.e. hyphae that lack septae) in which nuclei and cell organelles can freely move through the hyphal network (Marleau et al. 2011). Although, AMF have been shown to harbour a complete and functional meiotic and pheromone-sensing pathway, as well as sets of mating types (Halary et al. 2013; Halary et al. 2011; Pelin et al.,2012), sexual reproduction has not yet been observed. Limitations in distinguishing morphological characters and molecular traits are among the major factors limiting progress in AMF research. In addition, intra-isolate genetic polymorphism and nuclear organization are still under debate. Furthermore, many lines of evidence support the theory that AMF harbour genetically different nuclei in a common cytoplasm, making these organisms highly unusual (Angelard et al. 2010; Beaudet et al., 2015; Boon et al. 2015; Boon et al. 2010; Boon et al. 2013; Hijri and Sanders 2005; Kuhn et al. 2001). In contrast to the nuclear genomes of AMF, mitochondrial coding-genes such as *rnl* and *cox1,* as well as complete mtDNAs, have been reported to be homogeneous in many AMF taxa (Borstler et al., 2008a; Lee and Young, 2009; Raab et al., 2005; Thiery et al., 2010), although this has been challenged by Beaudet et al., (2015). Investigating large datasets of mtDNAs could offer opportunities to advance our understanding of the population genomics and phylogeny of AMF.

To date, 16 fully sequenced Glomeromycota mt genomes have been published, representing two major glomeromycotan families: the Glomeomeraceae (14 mtDNAs) and the Gigasporaceae (two mtDNAs) (Beaudet et al., 2013a; Beaudet et al., 2013b; de la Providencia et al., 2013;

Formey et al., 2012; Nadimi et al., 2012; Pelin et al., 2012; Nadimi et al., 2014). These mtDNA datasets offer opportunities to compare the intra- and inter-specific diversity of mt genomes and genes among closely related AMF taxa, as well as the usefulness of various mt genes in phylogenetic analyses. In this study, we sequenced and mapped the mtDNAs of two AMF species, *Rhizophagus irregularis* DAOM240159 and *Glomus aggregatum* DAOM240163, to augment the publically available set of AMF mtDNAs. We performed a mitogenomics comparative analysis, as well as principal coordinate and partitioning analyses, in order to: (1) assess the phylogenetic signals of mt coding genes; (2) evaluate their potential use in inferring AMF phylogenies; and (3) unravel evolutionary relationships within the *Rhizophagus* genus.

5.4. Material and methods

5.4.1. Fungal material and DNA extraction

Spores and mycelium of a *Glomus aggregatum* isolate (DAOM240163) and of a *Rhizophagus irregularis* isolate (DAOM240159) were cultivated *in vitro* on a minimal (M) medium with carrot roots that had been transformed with *Agrobacterium rhizogenes* (Bécard and Fortin 1988). The medium was dissolved using a solution of 0.82 mM sodium citrate and 0.18 mM citric acid, and fungal mycelia were collected by sieving the suspension. The collected fungal material was checked for root contamination under a binocular microscope. DNA extraction was performed using the DNeasy Plant kit (Qiagen, Mississauga, ON) according to the manufacturer's instructions.

5.4.2. Sequencing, assembly and gene annotation

Whole genome shotgun sequencing was performed using Roche 454 GS FLX Titanium technology, with one full run per isolate. Sequencing was performed at the Genome Quebec Innovation Center (McGill University, Montreal, QC). The reads were *de novo* assembled with Newbler v2.9 at the Genome Quebec facility, and mitochondrial contigs were identified and pooled for each isolate. Gene annotation was performed using an automated mt annotator MFannot (http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl), followed by manual inspection. The *R. irregularis* DAOM197198 mtDNA (accession number HQ189519) was used as reference for sequence analyses and mtDNA comparison. Sequencing of the PCR products (acquired for joining mtDNA contigs) was performed using Sanger technology at the Genome Quebec Innovation Center (McGill University, Montreal, Canada). The annotated mtDNAs of *R. irregularis* DAOM240159 and *G. aggregatum* DAOM240163 were deposited in GenBank under the accession numbers KM586389 and KM586390, and circular maps were built using OGDraw v. 1.2 software (Lohse et al. 2013). Comparative sequence analyses were performed using the National Center for Biotechnology Information (NCBI) genomic database and multiple sequence alignment. Nucleotide blast was used for similarity analyses, followed by amino acid comparisons for more distinct sequences. A search with tBLASTx was also performed using a minimum E-value cutoff of $1x10^{-10}$ and 50% minimum identity.

5.4.3. Polymerase Chain reactions (PCR)

As mentioned above, conventional PCRs were used to fill the gaps between mt contigs in order to complete the mtDNA genome for each isolate. Conventional PCRs were performed in a final volume of 50 µl containing 1X PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM of each

deoxynucleotide triphosphate (dNTP), 0.5 µM of each primer, 1 unit of Platinum *Taq* DNA Polymerase (Life Technologies, Burlington, ON), and 5 to 20 ng of genomic DNA was used as a template. Cycling parameters were: 94°C for 90 sec, followed by 38 cycles of 94°C for 1 min, annealing temperature (variable based on the designed primers but generally ~60°C) for 30 sec, 72°C for 90 sec and a final elongation at 72°C for 5 min. PCR products were directly sequenced as described above.

5.4.4. Datasets and sequence alignments

Sixteen completely sequenced mt-genomes whose length ranged between 59,633 bp (*G. cerebriforme*) and 97,349 bp (*Gigaspora rosea*) were used in comparative analyses. The complete mtDNA alignment was evaluated for phylogenetic and comparative mitogenomic analyses. 16 genes containing 14 protein-coding genes (*atp6, atp8, atp9, cox1, cox2, cox3, nad1, nad2, nad3, nad4, nad4l, nad5, nad6* and *cob*) and two ribosomal coding subunits (*rns* and *rnl*) were concatenated to construct the ‘‘supergene’’ set. Multiple sequence alignment was performed using ClustalW as implemented in MUSCLE version 3.8.31 (Edgar 2004).

5.4.5. Phylogenetic analyses

Phylogenetic analyses were performed for all mt genes from fully sequenced mtDNAs of AMF isolates (Table 5.1). Nucleotide alignments of all datasets were used to infer maximum likelihood (ML) phylogenies as implemented in RAxML (Stamatakis et al. 2008), using the RAxML-HPC2 on XSEDE web-server at CIPRES [http://www.phylo.org] (Miller et al.,2010).

Node support was determined using rapid bootstrapping with 1000 replicates. All datasets were analyzed using the GTRGAMMA model. Phylogenies were also generated for the datasets using concatenation of all mt genes (supergene set) and for a subset of genes, using concatenation of the four best-performing mt genes (genes that showed similar tree topology to that of supergene).

5.4.6. Statistical analysis of phylogenies

Topologies of each dataset, composed of each of the 16 mt genes, or a subset of the four best-performing mt genes with different combinations were compared to the supergene set. Statistical analyses of the comparisons of phylogenetic signals for different datasets, as well as comparisons of the topologies derived from mt genes and that of supergene, were performed using Shimodaira-Hasegawa (SH) tests, implemented in RAxML using '-f h' parameters locally. The SH test provides log-likelihood values and the variation scores (Ln *L*) between the supergene topology and each of the topologies inferred from single mt gene subsets, with respect to the standard deviation (SD). This test was performed for phylogenies inferred using the ML approach based on nucleotide datasets.

5.4.7. Distance matrix, principal coordinate analysis, and partitioning analysis

All analyses were conducted in R v.3.0.2 (R Foundation for Statistical Computing; available at http://www.R-project.org). After using the 'read.alignment' function of the 'seqinr' packages (Charif and Lobry 2007), distance matrices were computed using 'dist.hamming', 'dist.dna' and 'dist.ml' of the 'phangorn' package (Paradis et al. 2004; Schliep 2011). The similarity relationship between isolates was assessed with a principal coordinate analysis (PCoA) plot, using the 'pcoa' function of the 'ape' package (Paradis et al. 2004). In order to detect isolate

clusters within the plot, partitioning analysis were computed using a K-means clustering approach, and analyses were computed using 'kmeans' of the 'fpc' package (Hennig 2010), looking for a bend in the sum of squared roots (SSE). We also used partitioning around k-medoids to estimate the number of clusters, using the 'pamk' function of the 'fpc' package (Hennig 2010).

5.5. Results and Discussion

5.5.1. Overview of mtDNA diversity and evolution in the Glomeromycota

Fourteen complete mtDNAs belonging to two orders (Glomerales and Diversisporales) and three genera (*Glomus*, *Rhizophagus* and *Gigaspora*) were retrieved from public databases, and this dataset was augmented by two newly sequenced mtDNAs, resulting in a total of 16 mtDNAs from the Glomeromycota phylum (Table 5.1). Although mt genome structure and genome synteny varies somewhat among the studied isolates, they all contain 14 protein-coding genes and 2 ribosomal subunit genes. *R. irregularis* isolate DAOM240159 and *G. aggregatum* isolate DAOM240163 are the two newly sequenced mt genomes, measuring 72,293 bp and 69,505 bp, with G + C contents of 37.1% and 37.3%, respectively (Fig. 5.1). Both mitochondrial genomes harbor the full set of typical fungal genes, such as three subunits of ATP synthase (*atp6*, *atp8* and *atp9*), three cytochrome oxidase subunits (*cox1*, *cox2* and *cox3*), seven subunits of the NADH dehydrogenase (*nad1*, *nad2*, *nad3*, *nad4*, *nad5*, *nad6* and *nad4L*) and apocytochrome b (*cob*). They encode 25 tRNAs and small and large ribosomal subunits (mtSSU and mtLSU) (Fig. 5.1). In total, 41 genes encoded on the same strand in both genomes are found in each mtDNA (Fig. 5.1). The annotated sequences of *R. irregularis* DAOM240159 and *G. aggregatum*

DAOM240163 were deposited in GenBank under the accession numbers KM586389 and KM586390, respectively. Our phylogenetic analysis, based on the concatenation of all individual mt genes (supergene), clustered *G. aggregatum* within the *R. irregularis* group. However, using sequence comparison of intergenic regions, we were able to discriminate *G. aggregatum* and *R. irregularis* isolates (Table 5.2).

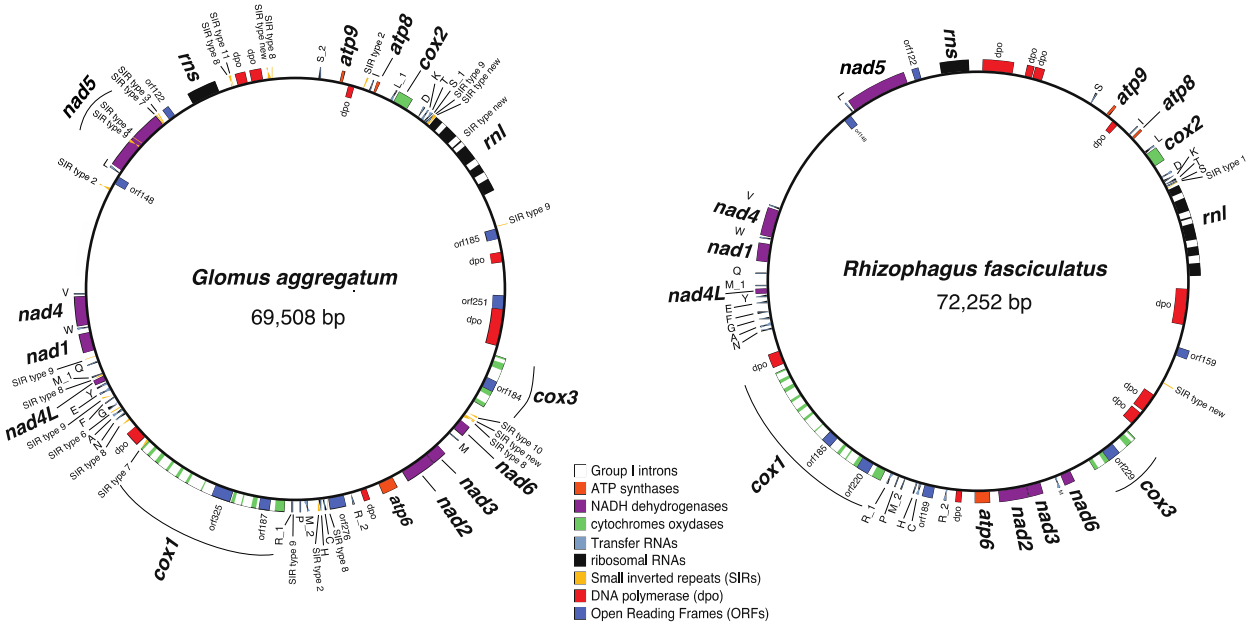


Figure 5.1. Comparison of the mitogenomes of the *Glomus aggregatum* isolate (DAOM240163) and the *Rhizophagus irregularis* isolate (DAOM240159).

The circular mapping genomes were opened upstream of *rnl*. Genes on the outer and inner circumference are transcribed in a clockwise and counterclockwise direction, respectively. Gene and corresponding product names are *cox1–3*, cytochrome c oxidase subunits, *atp6*, *8*, *9*, ATP synthase subunit*s*; cob, apocytochrome b; *nad1–4*, *4L*, *5–6*, NADH dehydrogenase subunits; *rnl*,

rns, large and small subunit rRNAs; A-W, tRNAs, the letter corresponding to the amino acid specified by the particular tRNA followed by their anticodon.

Table 5.1. Arbuscular mycorrhizal fungal isolates used in our mitogenomic analyses.

Species	solates	Cultures	Accession #	Length	Number of						Number of introns in PCGs				
					PCG	rRNA	tRNA	SIR			*cox*	*cox*			Reference
					s	s	s	s	dpos	*nad5*	*1*	*3*	*cob*	
R. irregularis	494	*In vitro*	FJ648425	70,606	14	2	26	NM	NM	2	11	3	4	Lee & Young 2009
R. irregularis	DAOM197198	*In vitro*	HQ189519	70,800	14	2	25	NM	NM	2	11	3	4	Nadimi et al. 2012
R. irregularis	DAOM234179	*In vitro*	KC164354	75,075	14	2	25	31	10	2	11	3	4	Beaudet et al. 2013b
R. irregularis	DAOM240415	*In vitro*	JX993113	70,781	14	2	25	NM	8	2	11	3	4	de la Providencia et al. 2013
R. irregularis	DAOM234328	*In vitro*	JX993114	68,994	14	2	25	NM	8	2	11	3	4	de la Providencia et al. 2013
R. irregularis	MULC46239	*In vitro*	JQ514223	87,755	14	2	25	NM	NM	2	12	3	NM	Formey et al. 2012
R. irregularis	MULC46240	*In vitro*	JQ514225	74,798	14	2	25	NM	NM	2	9	3	NM	Formey et al. 2012
R. irregularis	MULC43204	*In vitro*	JQ514224	87,755	14	2	25	NM	NM	2	12	3	NM	Formey et al. 2012
Rhizophagus sp.	DAOM229456	*In vitro*	JX065416	87,763	14	2	25	NM	2	2	14	3	5	Beaudet et al. 2013a
Rhizophagus sp.	DAOM240422	*In vitro*	KC164355	86,170	14	2	25	29	10	2	11	3	4	Beaudet et al. 2013b
Rhizophagus sp.	DAOM213198	*In vitro*	KF591215/6	91,020	14	2	25	20	11	3	13	3	7	Namidi et al. 2014 Submitted
R. irregularis	DAOM240159	*In vitro*	KM586389	72,293	14	2	25	2	9	2	10	3	4	Namidi et al. 2014 not published
Glomus aggregatum	DAOM240163	*In vitro*	KM586390	69,505	14	2	25	33	7	2	9	3	4	Namidi et al. 2014 not published
Glomus cerebriforme	DAOM227022	*In vitro*	KC164356	59,633	14	2	25	40	2	1	5	2	2	Beaudet et al. 2013b
Gigaspora rosea	DAOM194757	*In vivo*	NC_016985	97,350	14	2	25	NM	2	3	10	–	1	Nadimi et al. 2012
Gigaspora margarita	BEG34	*In vivo*	NC_016684	96,998	14	2	25	NM	NM	3	7	1	1	Pelin et al. 2012

Table 5.2. Sequence comparison of intergenic regions between *R. irregularis* 197198 and *G aggregatum*.

IGRs	Variable regions position
rnl-cox2	41 bp insertion in position 1042 in *G. aggregatum*
nad3-nad6	37 bp insertion in position 88 in *G. aggregatum*
nad5-cob	40 bp insertion in position 927 in *G. aggregatum*
nad4L-cox1	Polymorphism from 748-787bp in DAOM197198 and 745-795 in *G. aggregatum*
atp9-rns	65 bp deletion from position 5550 in *G. aggregatum* Polymorphism from 5410-5424 bp in DAOM197198 and 745-795 in *G. aggregatum*
nad6-cox3	36 bp insertion in position 557 in *G. aggregatum*
cox3-rnl	8628 bp insertion in position 1 in *G. aggregatum* About 7500 bp polymorphism at position 404 of DAOM197198 (9034 in *G. aggregatum*)

G. aggregatum was described by Schenck and Smith in 1982 after being found in the roots of citrus trees in Florida (Schenck and Smith 1982). Spores can be formed individually in the soil and in roots or aggregates without a peridium. These spores are globose, subglobose, obovate, and cylindrical to irregular. In fact, *G. aggregatum* and *R. irregularis* are very close according to the morphological characters of their spores, with the only tangible difference being that *G. aggregatum* spores possess an inner laminated wall that strongly reacts to Melzer's reagent, while the inside of spores are sometimes differentiated (Sokolski et al. 2011). A phosphate transporter gene was able to slightly discriminate *G. aggregatum* from a dozen *R. irregularis* isolates (Sokolski et al. 2011), while other genes such as ATPase H+, ATPaseF0F1, as well as ITS, 18R and 25S rRNA genes were not able to discriminate these two species. In a recent reclassification of the Glomeromycota, the position of *G. aggregatum* was classified as uncertain (Schussler et al. 2010).

5.5.2. Assessment of the phylogenetic signal of single mt genes and subsets of genes

Maximum likelihood phylogenies were inferred from nucleotide sequence alignments of the complete set of mt genes (16 coding genes in total). We assessed the phylogenetic signal of single mt genes, which was compared with the supergene signal via Shimodaira-Hasegawa (SH) test (Fig. 5.2). These comparisons allowed us to determine the performance of single mt genes, and to determine the best combination of gene subsets with phylogenies similar to the supergene phylogeny. Twelve topologies inferred from single mt genes showed a significant reduction in log-likelihood (Ln *L*) and statistically different topologies from that shown in the supergene topology (Fig. 5.2).

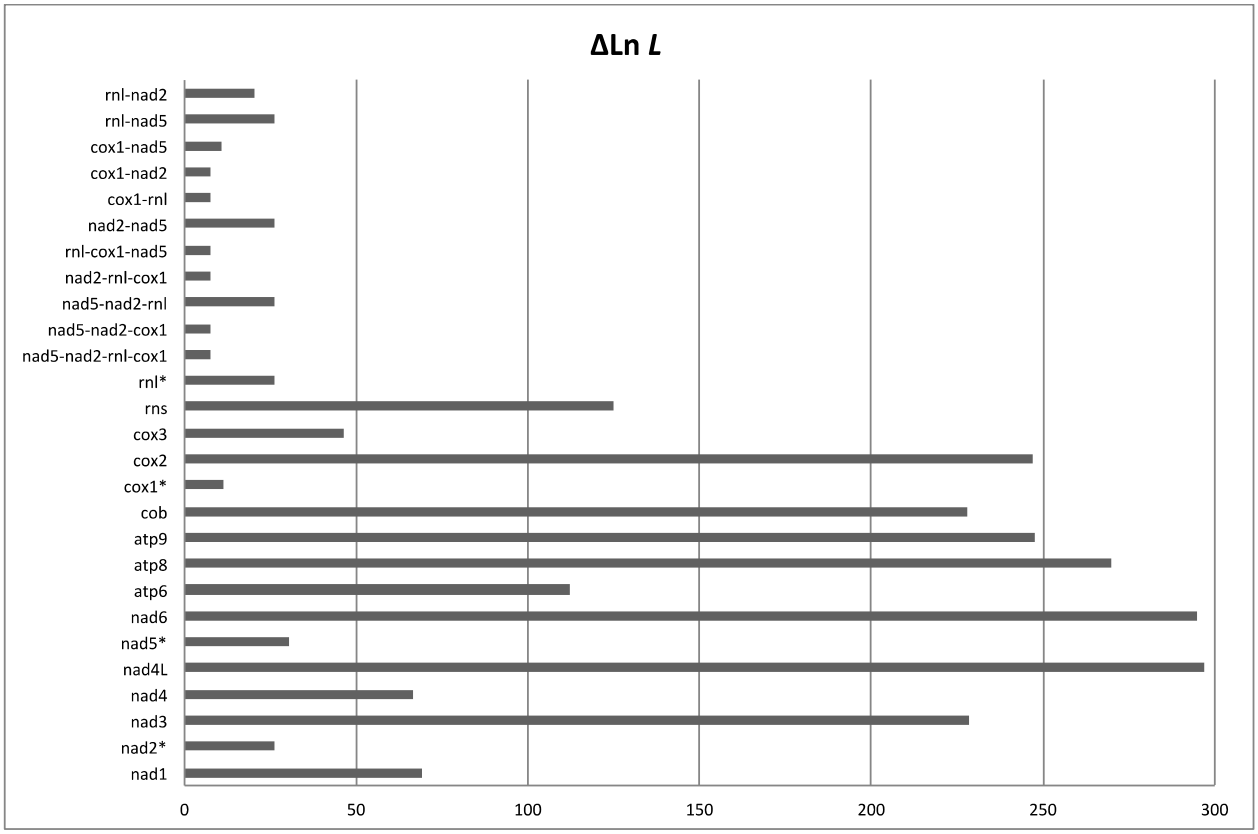


Figure 5.2. Difference in log-likelihood measured via SH test by comparing phylogenetic topologies inferred from single mt genes relative to that of the ‘‘supergene’’ set.

Difference in Ln *L* compared to the supergene phylogeny for all 16 mt genes (y-axis) has been shown on the x-axis as ΔLn *L*. The four best-performing genes (i.e., *cox1*, *rnl*, *nad5* and *nad2*) are indicated by *. Phylogenetic analyses were performed under ML and SH test implemented by RAxML locally.

However, the other four genes (i.e. *rnl*, *nad5*, *cox1* and *nad2*), which were the longest mt genes, reproduced similar topologies to the supergene tree, with much more subtle decreases in Ln *L* relative to the supergene phylogeny (Fig. 5.2). Visual inspection, coupled with statistical tests of single gene topologies, showed that the other mt genes provided poorly supported trees for two main reasons: (i) genes such as *atp9*, *cob*, *cox2* and *rns* could not resolve distant taxa such as *Gigaspora* spp. and *Glomus cerebriforme* from *Rhizophagus* spp.; (ii) *nad4L*, *nad6*, and *atp8* genes provided low resolution for resolving *Rhizophagus* sp. isolates from other *Rhizophagus irregularis*, which has been showed to be divergent (e.g. topologies inferred from *cob* and *nad6* are demonstrated in Figure 5.3). Furthermore, we evaluated the minimum number of concatenated mt genes required to statistically reproduce a topology similar to that of the supergene using ML. A SH test of the combination of 'best-performing' genes revealed low decreases in log-likelihood scores (Fig. 5.2). Phylogenies inferred from combinations of genes containing *cox1* with *nad2* or *rnl*, provided topologies that are statistically identical to the supergene topology (Fig. 5.3), with the smallest decreases in Ln *L* (Fig. 5.2).

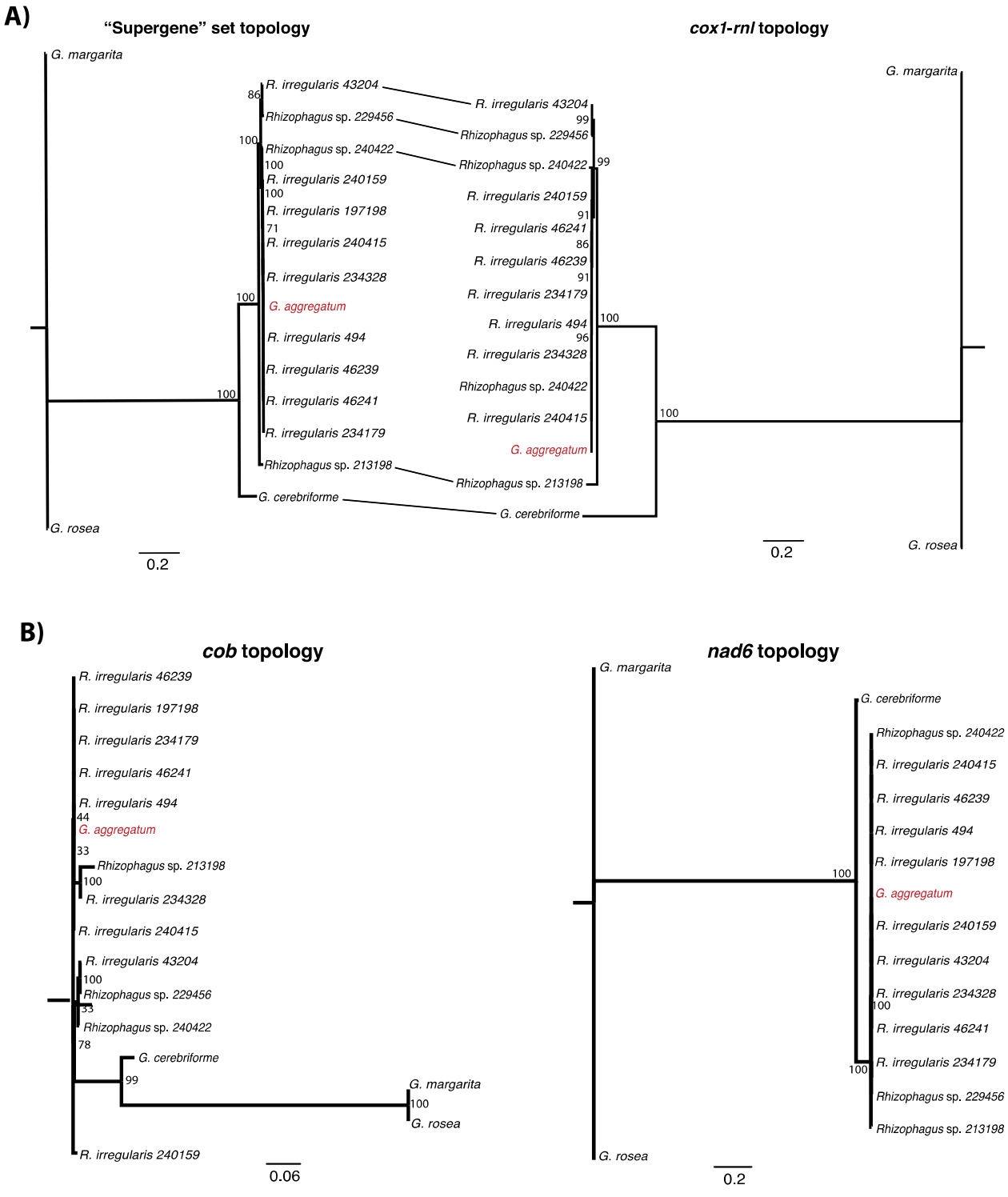


Figure 5.3. Topologies inferred from supergene set and concatenated set of *rnl* and *cox1* revealing their high concordance in addition to topologies inferred from *nad6* and *cob* as an example of individual genes with significant inconcordance to supergene topology.

A) The two phylogenies are significantly congruent based on statistical analyses of Shimodaira-Hasegawa (SH) with a difference of 7.6 in their Ln *L* values. Combination of *cox1* and *nad2* also showed a statistically congruent topology, which is not repeated here. The supergene set consists of a concatenation of all mt coding genes (i.e. 16 genes). *rnl*, *cox1* are the two largest mt genes. *G. aggregatum* is colored as red to highlight its position within the *Rhizophagus* clade. Numbers at nodes represent bootstrap support values. Phylogenies were inferred using DNA sequences and maximum likelihood (scale bars indicate replacements per site). B) Topologies inferred by *nad6* and *cob* DAN sequences genes showing significant decreases in Ln *L* values compared with supergene topology (ΔLn *L* of 294.6 and 228.1 respectively). *Cob* is not able to resolve distant taxa such as *Gigaspora* spp. and *Glomus cerebriforme* from *Rhizophagus* spp. while *nad6* possesses low resolution for resolving *Rhizophagus* sp. isolates from other *Rhizophagus irregularis*.

Here we demonstrate that choosing a single mt gene in phylogenetic analysis may not reflect the phylogeny of the complete set of mt coding genes. Some studies have shown that phylogenies from combinations of all mt genes or from complete mitogenomes represent the highest phylogenetic performance and more robust results than those produced by single genes and gene subsets (Duchene et al. 2011; Havird and Santos 2014). These studies identified the most informative mt regions (genes) and evaluated the minimum amount of data required to infer phylogenies that were similar to supergene or complete mtDNA topologies. Therefore, the most informative mt genes or regions might vary among taxonomic groups and organisms (Duchene et al. 2011; Havird and Santos 2014).

Mt genes *rnl* and *cox1* have previously been used independently to infer phylogenies among taxa of the Glomeromycota (Borriello et al. 2014; Borstler et al. 2008b; Raab et al. 2005). Our study demonstrates that phylogenetic analysis that combines *cox1* and either *rnl* or *nad2* performs better than analysis on individual genes, not only considering the performed phylogenetic analyses, but also the evolutionary features of each individual gene. There are certain

evolutionary aspects of *rnl* and *cox1* that need to be considered prior to their use in phylogenetic analyses of either organelles or organisms. For example, it is known that *rnl* is able to carry inserts and introns in variable regions that are not excised at the RNA level (Raab et al. 2005). Thus, it is often not easy to align mitochondrial rDNA sequences across species, due to introns, insertions, and the loss of domains. There is still no reliable method for inferring precise ends of the *rnl* gene, as the sequences of both terminal regions are highly variable. The only method for precise annotation is RNA analysis (Lang et al. 1987; Nadimi et al. 2012). Therefore, *rnl* can be useful for short distance species phylogenies, but would be problematic for phylogenies that attempt to resolve among fungal clades or deeper levels. There are also several reports regarding the evolutionary history and structure of *cox1*, which hampered the use of this gene for phylogenetic analyses and barcoding. For example, Nadimi et al. (2012) reported the occurrence of a group I intron-mediated trans-splicing in *cox1* of *Gigaspora rosea* that has been previously described in several other organisms (Burger et al. 2009; Grewe et al. 2009; Hecht et al. 2011; Nadimi et al. 2012; Pombert and Keeling 2010). Frequent horizontal gene transfer within the introns of this gene, which has been reported in the literature (Beaudet et al. 2013a; Lang and Hijri 2009), may also complicate the sequencing and assembly of this region, as well as downstream phylogenetic analyses. The above-mentioned findings limit the usefulness of individual *cox1* and *rnl* genes for phylogenetic analyses of AMF. Combinations of multiple genes (e.g., different combinations of *cox1*, *rnl* and *nad2*) may help to overcome these drawbacks.

5.5.3. Assessment of the relationship between isolates

To further evaluate the relationships between all 16 isolates used in this study, we performed a principal coordinate analysis (PCoA) based on the distance matrix obtained from mitochondrial

genomic alignments (Fig. 5.4). The PCoA plot shows that all *Rhizophagus* spp. isolates and *G. aggregatum* group together. Of the other three isolates, *G. cerebriforme* is apart, while both *Gigaspora spp.* cluster more closely together. Both partitioning methods used (k-means and k-medoids) created two groups, corresponding to the two orders (Glomerales and Diversisporales) used in the study. When inferring three groups instead of two on the plot, only *G. cerebriforme* separates from the Glomerales group to form the third group. To further assess the cluster of *Rhizophagus* isolates, another principal coordinate analysis (PCoA) was computed, excluding both *Gigaspora* species and *G. cerebriforme* (Fig. 5.4). Both partitioning methods inferred three groups, the first including solely *Rhizophagus* sp. DAOM213198, the second including *R. irregularis* MUCL43204 and *Rhizophagus* sp. DAOM229456, and the third including the remaining nine *Rhizophagus* isolates and *G. aggregatum*. One *Rhizophagus* sp. isolate, DAOM240422, clustered with the other nine *Rhizophagus* isolates, but with a weaker similarity (Fig. 5.4). These results support the *G. cerebriforme* isolate as genetically distant from all *Rhizophagus* isolates and from the *G. aggregatum* isolate used in this study. As discussed above, spore morphology and phosphate transporter phylogeny are able to marginally discriminate between *G. aggregatum* from *Rhizophagus* isolates. However, our analysis based on the complete mt genome shows that other *Rhizophagus* isolates, such as *R. irregularis* MUCL43204 and *Rhizophagus* sp. DAOM213198/DAOM299456 are genetically more distant from the other nine *R. irregularis* isolates than is *G. aggregatum*. This evidence supports the position of *G. aggregatum* within the *Rhizophagus irregularis* 'species complex'.

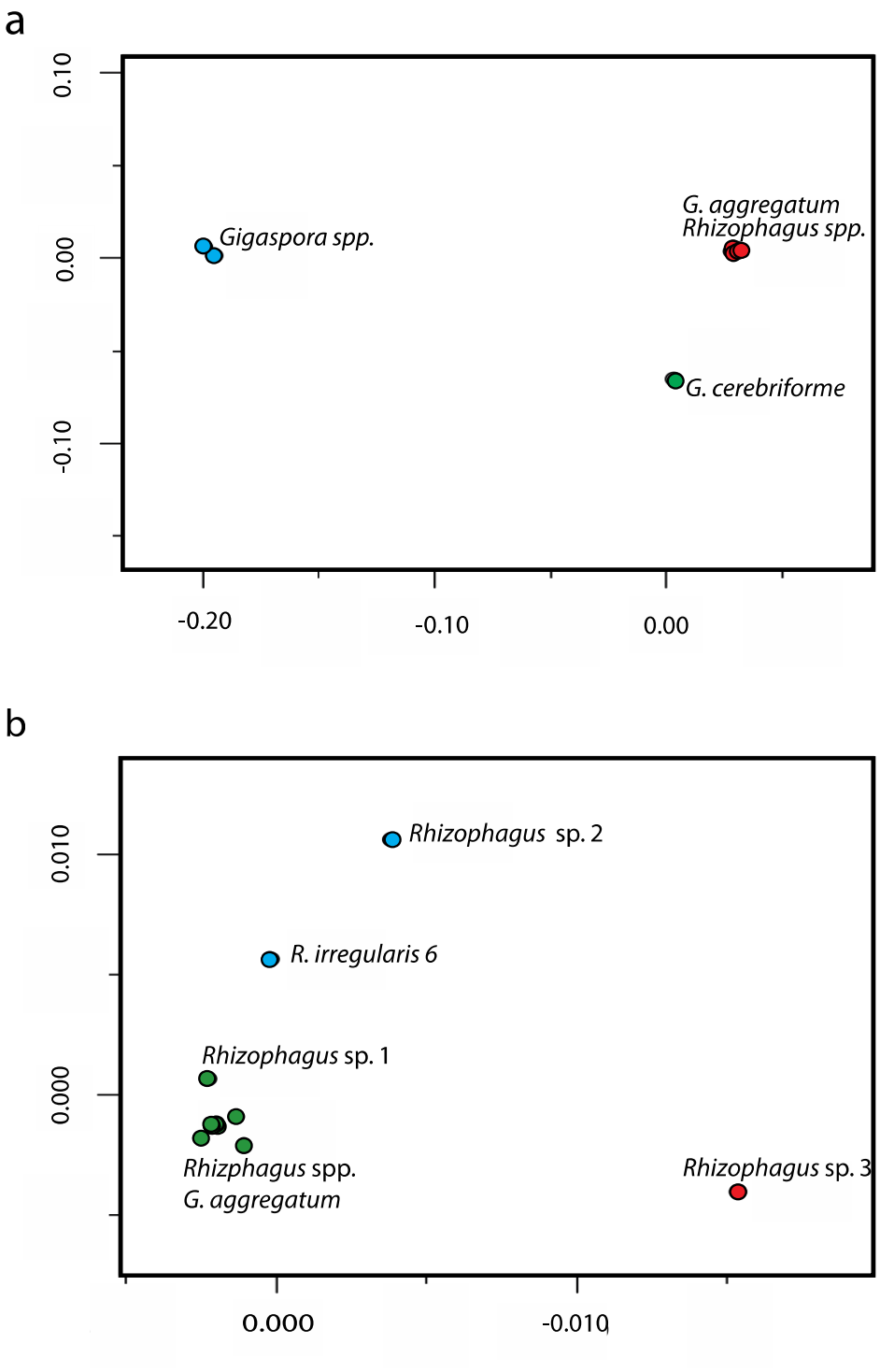


Figure 5.4. PCoA plot based on the distance matrix obtained from mitochondrial genomic alignments.

(**a**) using all 16 mtDNAs examined in this study. *Gigaspora* spp. (*G. rosea,* and *G. margarita*) and *Glomus cerebriforme* are distinguished from the *Rhizophagus* spp. and *G. aggregatum*

isolates. Two groups inferred by clustering analysis are shown in colors **(b**) Using 12 isolates of *Rhizophagus* spp. and *G. aggregatum*. The three groups inferred by clustering analysis are shown in different colors.

5.5.4. Unraveling the *Rhizophagus irregularis* complex

Comparative mitogenomics analysis of *R. irregularis* isolates (Fig. 5.5) and *Rhizophagus* spp. reveals a high level of homogeneity within the coding sequences, while intronic and intergenic regions showed substantial variable regions, which is consistent with the rate of evolutionary constraint based on the functions of different mtDNA regions. However, comparative analyses using ML and PCoA of coding regions suggest the occurrence of unorthodox recombinations (Fig. 5.6) among some *Rhizophagus* spp. isolates (DAOM229456, DAOM240422 and DAOM213198) with uncertain classification, grouping them apart from other known *R. irregularis* isolates. Genetic recombinations may have taken place at different points during the evolutionary history of the different isolates or of a recent common ancestor, although horizontal gene transfer of recombined regions is also a possibility. A shared common ancestor seems to be the most parsimonious and plausible, due to similar signatures observed in recombined regions (Fig. 5.6). Mt genes *atp6*, *cox2*, *nad3*, *cob*, *nad4* and *atp9* are examples of these recombined mt genes within *Rhizophagus* spp. (Beaudet et al. 2013a; Beaudet et al. 2013b).

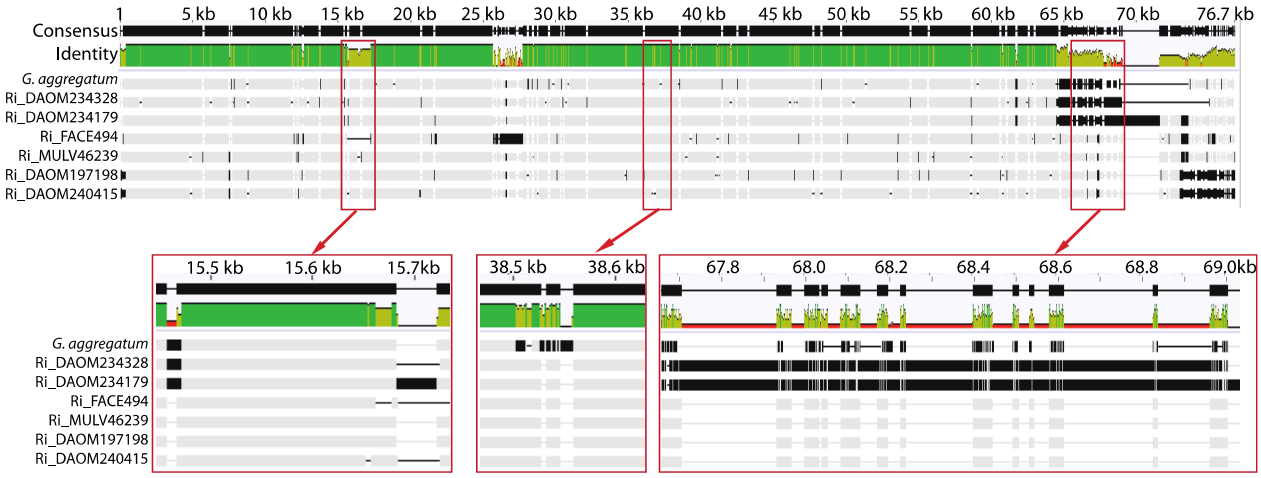
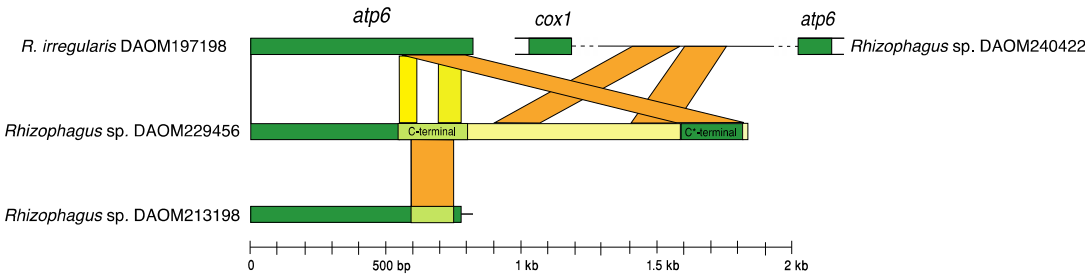
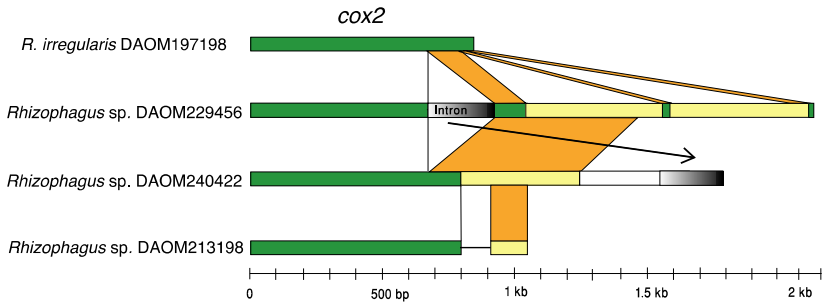
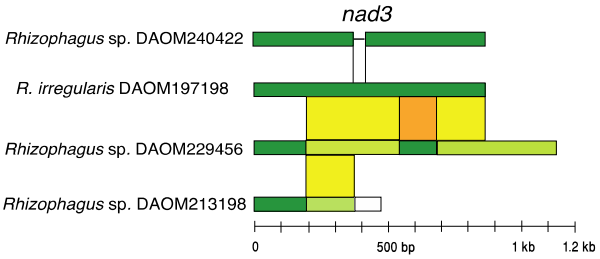


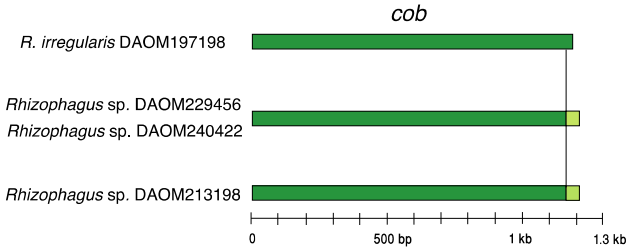
Figure 5.5. Complete mtDNA alignment of *R. irreglaris* isolates and *G. aggregatum* showing sequence variations in intergenic regions, while coding genes are almost identical.

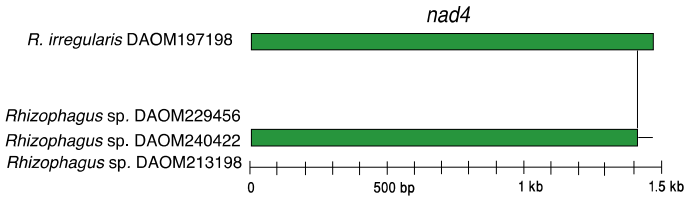
Comparative analyses of the intergenic regions of *Rhizophagus* spp. showed relatively higher levels of similarity in up- and down-stream sequences of coding genes compared to other intergenic regions. Up- and down-stream sequences of coding genes include 5′ and 3′ extensions of genes (known as 5′ and 3′ untranslated regions), which are transcribed but not translated. The conservation of such regions is due to their function, therefore resulting in a slower rate of evolution relative to non-functional regions. Consequently, comparative analyses of multiple sequence alignments of complete mtDNAs grouped *Rhizophagus* spp. isolates among the *Rhizophagus irregularis* species, which we suggest to be considered as a species complex (e.g. *Rhizophagus irregularis* complex).


atp6
cox1
atp6
R. irregularis DAOM197198
Rhizophagus sp. DAOM240422
Rhizophagus sp. DAOM229456
C-terminal
C*-terminal
Rhizophagus sp. DAOM213198
0
500 bp
1 kb
1.5 kb
2 kb


cox2
R. irregularis DAOM197198
Rhizophagus sp. DAOM229456
Intron
Rhizophagus sp. DAOM240422
Rhizophagus sp. DAOM213198
0
500 bp
1 kb
1.5 kb
2 kb


nad3
Rhizophagus sp. DAOM240422
R. irregularis DAOM197198
Rhizophagus sp. DAOM229456
Rhizophagus sp. DAOM213198
0
500 bp
1 kb
1.2 kb


cob
R. irregularis DAOM197198
Rhizophagus sp. DAOM229456
Rhizophagus sp. DAOM240422
Rhizophagus sp. DAOM213198
0
500 bp
1 kb
1.3 kb


nad4
R. irregularis DAOM197198
Rhizophagus sp. DAOM229456
Rhizophagus sp. DAOM240422
Rhizophagus sp. DAOM213198
0
500 bp
1 kb
1.5 kb

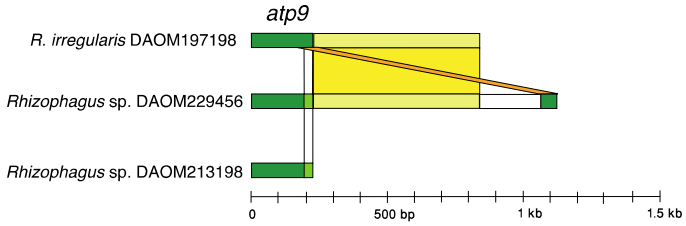

atp9
R. irregularis DAOM197198
Rhizophagus sp. DAOM229456
Rhizophagus sp. DAOM213198
0
500 bp
1 kb
1.5 kb

Figure 5.6. Linear genome representations of some mt genes, harbouring the relative recombinations within the *Rhizophagus* sp. complex to compare with *R. irregularis*.
Corresponding genes (*atp6*, *cox2*, *nad3*, *cob*, *nad4* and *atp9*) among *R. irregularis* DAOM197198 and *Rhizophagus* sp. DAOM240422, DAOM213198 and DAOM229456 are linearized and compared using projections. Projections in yellow represent nucleotide identity at more than 90%, while projections in orange illustrate the identity of more that 90% at the amino acid level. Projections with no color were used to facilitate detection of related regions. The intron formed in *Rhizophagus* sp. DAOM229456 is boxed in gray, which is transferred to the intergenic region after *cox2* in *Rhizophagus* sp. DAOM240422. Nucleotide and amino acid identity comparisons were performed via blastn and tblastx among the studied AMF isolates.

5.6. Conclusion

We implemented the existing mtDNA dataset with two newly completed and annotated mt genomes. Our mitogenomic comparative analysis among several AMF species and isolates allowed us to assess the phylogenetic signals of individual mt genes in comparison to the supergene set. A statistical SH test revealed that the evolutionary signal differed significantly between single mt genes, subsets of genes, and the supergene in the Glomeromycota. The combination of *cox1* with either *rnl* or *nad2* genes is a promising subset of mt genes that could be useful in inferring phylogenies within Glomeromycota. However, implementation of the mtDNA dataset with members of other families and genera is needed to cover the diversity of the Glomeromycota phylum and to infer robust phylogenies. Further studies on the comparison and potential combination of mt gene subsets, supergenes, and nuclear ribosomal rRNA genes could provide a higher phylogenetic resolution, and may helpt to discriminate between closely related members of the Glomeromycota.

5.7. Acknowledgment

This work was supported by the *Natural Sciences and Engineering Research Council* of Canada (Discovery grant 328098-2012 to M.H.). Our thanks to Dr. B.F. Lang for bioinformatics assistance and access to an automated organelle genome annotation software, to Dr. Y. Dalpé for kindly providing fungal material, Dr. S. Joly for assistance in phylogenetic and statistical analyses, and Dr. T. Bell for commenting on the manuscript and English editing.

Presentation of article 5

The presence of a mixture of more than one type of mt genome results in heteroplasmy, which has never been observed in AMF. Typically in sexual organisms, biparental inheritance (paternal leakage) and/or occurrence of mutation result in coexistence of different mt haplotypes within a single cell. In case of AMF as asexual organisms, exchange of genomic material via a process called anastomosis raises question about the presence of a heteroplasmic state in them. Nonself recognition between members of the *Rhizophagus* genus has been documented (Croll et al. 2009; Angelard et al. 2010; Colard et al. 2011, Purin and Morton 2012) and has resulted in nuclear transmition to the progenies but exchange of divergent mitochondrial haplotypes has yet to be shown. In yeast, heteroplasmic state has been reported and it has ben shown that this condition is temporary due to different mechanisms such as segregation, leading to permanent homoplasmy (White and Lilley 1997; MacAlpine et al. 1998; Birky 2001).

In this study, we employed three genetically well-characterized *R. irregularis* isolates (DAOM197198, DAOM234328 and DAOM240415) to upgrade our knowledge about AMF mitogenome and creditability of it for AMF identification and to address the question, whether heteroplasmy occurs in AMF. These isolates were selected because they originate from distant geographical locations. In addition, de novo genome sequencing, assembly and annotation were performed for two other *R. irregularis* isolates, DAOM234328 (Finland) and DAOM240415 (Manitoba, Canada), and their results were compared with published mtDNAs (Lee and Young 2009; Formey et al. 2012; Nadimi et al. 2012). Subsequently, specific size-variable markers were designed for each isolate to detect mtDNA mixture after anastomosis between the strains in microcosm experiments.

The **Article 5** was published in New Phytologist. I assembled and annotated both of mtDNAs. It took 9 months to design and set up the pre-symbiotic and symbiotic experiments examining different methods to apply the most appropriate ones. By taking maternity leave, Ivan de La Providencia went through performing the experiments with assistant of Gabriela Rodriguez and afterwards he completed writing of the article. I also contributed to writing the first draft of the manuscript.

Chapter 6 - Detection of a transient mitochondrial DNA heteroplasmy in the progeny of crossed genetically divergent isolates of arbuscular mycorrhizal fungi

Ivan Enrique de la Providencia, Maryam Nadimi, Denis Beaudet, Gabriela Rodriguez Morales and Mohamed Hijri

Université de Montréal, Département de sciences biologiques, Institut de recherche en biologie végétale (IRBV), 4101 rue Sherbrooke Est, Montréal, QC, H1X 2B2, Canada.

Published in New Phytologist, 2013 Oct. 200(1): 211-21, doi: 10.1111/nph.12372.

6.1 Summary

- Nonself fusion and nuclear genetic exchange have been documented in arbuscular mycorrhizal fungi (AMF), particularly in *Rhizophagus irregularis*. However, mitochondrial transmission accompanying nonself fusion of genetically divergent isolates remains unknown.
- Here, we tested the hypothesis that mitochondrial DNA (mtDNA) heteroplasmy occurs in the progeny of spores, obtained by crossing genetically divergent mtDNAs in *R. irregularis* isolates.
- Three isolates of geographically distant locations were used to investigate nonself fusions and mtDNA transmission to the progeny. We sequenced two additional mtDNAs of two *R. irregularis* isolates and developed isolate-specific size-variable markers in intergenic regions of these isolates and those of DAOM-197198. We achieved three crossing combinations in pre-symbiotic and symbiotic phases. Progeny spores per crossing combination were genotyped using isolate-specific markers.
- We found evidence that nonself recognition occurs between isolates originating from different continents both in pre-symbiotic and symbiotic phases. Genotyping patterns of individual spores from the progeny clearly showed the presence of markers of the two parental mtDNA haplotypes. Our results demonstrate that mtDNA heteroplasmy occurs in the progeny of the crossed isolates. However, this heteroplasmy appears to be a transient stage because all the live progeny spores that were able to germinate showed only one mtDNA haplotype.

6.2 Keywords

Anastomosis; arbuscular mycorrhizal fungi (AMF); heteroplasmy; mitochondrial genome; nonself fusion; specific molecular marker

6.3 Introduction

The Glomeromycota phylum encompasses an ancient group of obligate symbiotic fungi colonizing c. 80% of plants to form arbuscular mycorrhizae (Smith and Read 2008). Long considered asexual organisms, this paradigm has recently been challenged by the discovery of intact meiotic machinery genes that are putatively functional (Halary et al. 2011), although apparent reproductive structures have yet to be observed. Arbuscular mycorrhizal fungi (AMF) reproduce clonally through large and multinucleated spores, and due to their coenocytic nature (i.e. living hyphae lacking septa and thereby nuclei, mitochondria and other organelles sharing a common cytoplasm), spores and mycelium can host communities of hundreds to thousands of nuclei showing substantial genetic variability (Kuhn et al. 2001; Hijri and Sanders 2005; Jany and Pawlowska 2010; Marleau et al. 2011) which persist at the transcriptional level (Boon et al. 2010).

Several studies have shown the impact of AMF on plant growth and health (Jeffries et al. 2003, Smith and Read 2008, van der Heijden et al. 2008, Ismail and Hijri 2012) soil quality and structure (Rillig and Mummey 2006), and ecosystem functioning and biodiversity (van der Heijden et al. 2006). When associated with plants, they often extend beyond roots into the surrounding environment to form an intermingled coenocytic and multinucleated extraradical mycelium network (Leake et al. 2004) that facilitates up to 90% of phosphorus (P) plant-uptake and also enhances nitrogen (N)-acquisition under natural conditions (reviewed in Bolduc and Hijri (2011)). AMF structures can also link plants from the same or different species and/or families via the formation of hyphal fusions (i.e. anastomosis) (Giovannetti et al., 2004), creating

an extensive common mycorrhizal network. As previously demonstrated by Giovannetti et al. (1999), anastomosis also plays a key role in genetic exchange, either within the same isolate or between closely related isolates that are genetically divergent (Croll et al. 2009). This mechanism is likely one of the most important factors contributing to the maintenance of genomic composition among members of the Glomeromycota (Bever and Wang 2005, Corradi et al. 2007, Croll et al. 2009; Angelard and Sanders 2011; Colard et al. 2011). However, such huge genetic diversity within AMF populations (Corradi et al. 2007; Bonen 2008) or even in successive monosporal subcultures (Ehinger et al. 2012) makes it difficult to use only nuclear markers to address questions related to AMF community structure, diversity and function. In contrast to nuclear genes, mitochondrial genomes appear to be homogeneous within AMF isolates.

This was clearly demonstrated by sequencing the first complete mitochondrial genome (Lee and Young 2009) of the *Rhizophagus irregularis* isolate 494 (Błaszk., Wubet, Renker & Buscot) C. Walker & A. Schussler – syn. *Glomus irregulare* Błaszk., Wubet, Renker & Buscot (Blaszkowski et al. 2008; Stockinger et al. 2009; Schüßler and Walker 2010), followed by many other genomes such as *Gigaspora margarita*, *G. rosea*, four isolates of *R. irregularis* and a closely related *Rhizophagus* species (Formey et al. 2012; Nadimi et al. 2012; Pelin et al. 2012; Beaudet et al. 2013). These genomes offer exceptional opportunities for designing specific markers to identify AMF strains either in natural populations or in mesocosm/microscosm based-experiments (Lang and Hijri 2009; Formey et al. 2012; Beaudet et al. 2013). It was previously assumed that, based on incompatibility and nonself recognition mechanisms, anastomosis cannot occur between genetically distinct isolates (Giovannetti et al. 2003). However, many studies have challenged this point of view and have provided convincing contrary evidence (Croll and Sanders 2009; Angelard et al. 2010; Angelard and Sanders 2011). These studies used crossing experiments and specific genetic markers to show that nuclear material from both parents is

passed on to their progeny, thus demonstrating genetic exchange and segregation between genetically divergent isolates of the AMF *R. irregularis*. In addition, a recent study has clearly demonstrated that genetically different isolates of *Rhizophagus clarus* (synonym, *Glomus clarum*) originating from the same habitat and sampled close to each other underwent anatomosis (Purin and Morton 2012). This once again raises the questions of whether mitochondria are also exchanged via anastomosis between isolates of *R. irregularis* of different geographical origin, that is, heteroplasmy (i.e. the mixture of genetically different mtDNAs in a common cytoplasm) and whether this state is transient or is maintained into the progeny. Heteroplasmy is not rare in the fungal kingdom (Lesemann 2006) and has been observed in several eukaryotes (Barr et al. 2005; White et al. 2008). As suggested by researchers working on natural populations of the plant pathogen *Armillaria sp.* (Smith et al. 1990), possessing multiple types of mtDNA might confer an advantage under changing environmental conditions. This hypothesis has already been demonstrated for the model fungus *Saccharomyces cerevisiae* (Taylor et al. 2002; Hnatova et al. 2003) and the plant pathogen *Podosphaera leucotricha* (Lesemann 2006).

Heteroplasmy has never been studied in AMF, presumably due to the lack of reliable and appropriate mitochondrial markers (Lang and Hijri 2009; Lee and Young 2009; Formey et al. 2012). Consequently, we employed three genetically well-characterized *R. irregularis* isolates (DAOM-197198, DAOM-234328 and DAOM-240415) to address this question, which is an important one for advancing our knowledge of genetic exchange between AMF. These isolates were selected because they originate from distant geographical locations. In addition, *de novo* genome sequencing, assembly and annotation were performed for two other *R. irregularis* isolates, DAOM-234328 (Finland) and DAOM-240415 (Manitoba, Canada), and their results were compared with published mtDNAs (Lee and Young 2009; Formey et al. 2012; Nadimi et al. 2012). Subsequently, specific size-variable markers were designed for each isolate to detect

mtDNA mixture after anastomosis between the strains in microcosm experiments.

6.4 Materials and Methods

6.4.1 Growth conditions and maintenance of fungal cultures and roots

Monoxenically produced spores of *Rhizophagus irregularis*; DAOM-197198 (Pont-Rouge, Quebec, Canada) DAOM-234328 (Finland) and DAOM-240415 (Dufrost, Manitoba, Canada) were provided by the DAOM collection (Ottawa, Ontario, Canada). These three isolates were selected because they have different geographical origins and their mitochondrial genomes have been fully sequenced. Spores were subcultured in association with Ri T-DNA transformed chicory (Cichorium intybus) roots on a modified minimal (MM) medium (Becard and Fortin 1988) solidified with 0.4% (w/v) gellan gum (Sigma). Plates were incubated in the dark in an inverted position at 25°C. Several thousand spores and extraradical mycelia were obtained in a period of 12 wk. Ri T-DNA transformed chicory roots were routinely propagated by placing actively growing root apexes on MM medium and subsequent incubation at 25°C in the dark.

6.4.2 DNA extraction and sequencing

For DNA extraction, the extracted fungal spores and mycelia from the MM medium were dissolved in buffer citrate (pH 6.0) (Doner and Becard 1991) and washed with sterile water. Extracted fungal material from the MM medium was observed under a binocular microscope in order to detect and remove any root contaminants. DNA extraction was carried out using the DNeasy Plant Mini Kit (Qiagen), following the manufacturer's instructions. Whole genome shotgun sequencing was performed at McGill University's Genome Quebec Innovation Centre

using a Roche GS-FLX-Titanium sequencing platform (454 Life Sciences, Branford, CT, USA).

6.4.3 *de novo* assembly and sequence analysis

The total DNA of *R. irregularis* isolates DAOM-240415 and DAOM-234328 was sequenced and the resulting reads (257 880 and 238826, respectively) were assembled using Newbler (Genome Quebec Innovation Centre, McGill University, Montreal, PQ, Canada). Gene annotation was performed with MFannot (http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl), followed by manual inspection and notation of missing gene features. More information on MFannot is described in Nadimi et al. (2012). The annotated mitochondrial genomes of *R. irregularis* isolates DAOM-240415 and DAOM-234328 were deposited in GenBank under the respective accession numbers JX993113 and JX993114.

6.4.4 Molecular marker development

In order to develop isolate size-specific molecular markers for the mtDNA of three different *R. irregularis* isolates, the variable intergenic regions of their complete sequenced mitochondrial genomes were used for multiple alignments. Consequently, we developed specific primers for each isolate, which resulted in JX993114 three pairs of primers. In addition, these primers produced PCR fragments of different lengths (Table 6.1). All primers were tested for hairpin and dimer formation using AmplifX software and a SIGMA DNA calculator (http://www.sigma-genosys.com/calc/ DNACalc.asp). The specificity of these markers was achieved using PCR on amplified DNA of each isolate (Fig. 6.1). These markers made it possible to trace parental mtDNA haplotypes in crossing experiments.

Table 6.1. Isolate-specific primers used to discriminate the three *Rhizophagus irregularis* isolates

Primers	Codes	Primer sequences (5′- 3′)	Size (bp)	Position of primers in mtDNA		
				DAOM-197198	DAOM-240415	DAOM-234328
197198F	A	AGCAAATCTAAGTTCCTCAGAG	120	69523–69544		
197198R		TCCTGCCCCAGATACTCCAG		69624–69643		
240415F	B	AGAAGAGTTAATAAGTTCAACTGT	147		68561–68584	
240415R		AGACTAATAATACGAGATGCAG			68687–68708	
234328F	C	AAGAGGGGTGCATTCCGAAG	209			67641–67660
234328R		AGCTACGCTTTGCCTTAGCA				67831–67850

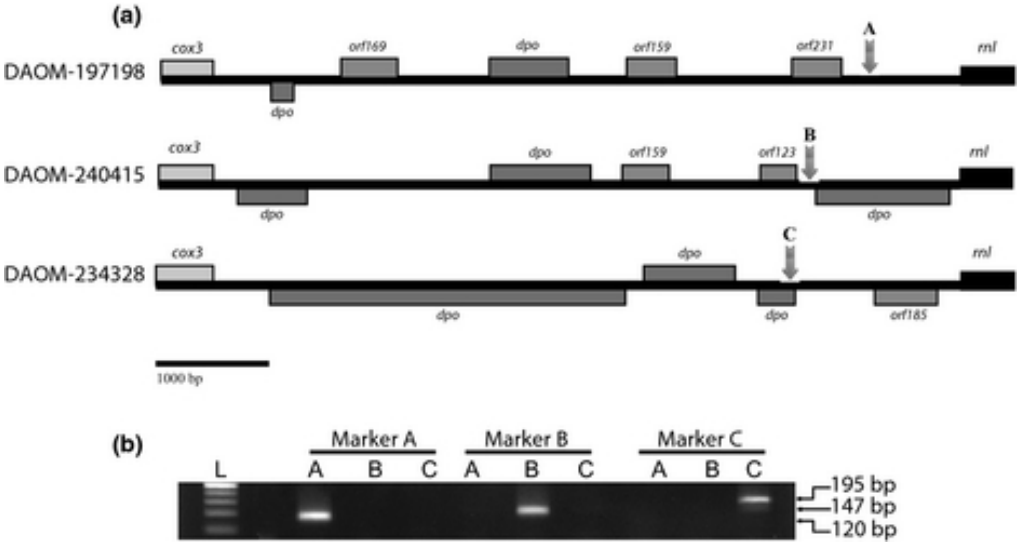


Figure 6.1. Isolate-specific mtDNA markers.

(a) Comparative view of the *cox3-rnl* intergenic regions of the three *Rhizophagus irregularis* isolates where primers were designed (arrows). Upper boxes (open reading frames, ORF) show direct orientation while bottom boxes are opposite orientation. (b) Gel electrophoresis showing the specificity of length-specific molecular markers developed for the *R. irregularis* isolates DAOM-197198 (A), DAOM-240415 (B) and DAOM-234328 (C). L, ladder (molecular marker).

6.4.5 Experimental set up

We set up two experiments in pre-symbiotic and symbiotic phases, respectively, to study the occurrence of anastomosis and subsequent exchange of mitochondria among the three isolates in crossing experiments. We also performed a single symbiotic experiment to evaluate the inheritance pattern of mitochondria from monosporal culture lines obtained from spores harvested from the interaction zone. Hyphal interactions were classified into four categories as described in Croll et al. (2009): (1) perfect fusion meant that hyphae entered into contact, fused and protoplasmic flow was established following anastomosis. No sign of incompatibility reaction was observed. (2) Pre-fusion incompatibility was characterized by profuse mycelium ramification at the elongation zone of the hypha that could be followed by septa formation at the hyphal apex, thus preventing contact between the approaching hyphae. (3) No interaction means that two hyphae (regardless of the type of interaction) did not show signs of rejection (i.e. wall thickening at the apex region, septa formation or change of growth direction) but intermingled (i.e. hyphae from different isolates or the same isolate grew in different directions and crossed, but did not touch each other) or overlapped (i.e. hyphae from different isolates or the same isolate entered into contact and continued to grow over each other). (4) Post-fusion incompatibility means that following homing, contact and fusion, the protoplasm of one hypha withdrew and a septum was formed between the two fused hyphae. Percentage of anastomosis for the categories perfect fusion and nonself fusion incompatibility was calculated by dividing the number of fusions by the total number of observed hyphal contacts (Purin and Morton 2012).

6.4.6 Pre-symbiotic experiment

Spores of the three isolates were extracted from the gel by solubilisation in buffer citrate (pH 6) (Doner and Becard 1991). A mono-compartment plate (35 9 10 mm) filled with 5 ml of MM medium was divided into four equal compartments. Before inoculation a colour-coded circle was marked at the bottom of the plate on each compartment and used to identify clusters belonging to the same or different isolates. On each compartment, ten clusters (a cluster was defined in this study as a group of spores bearing the same mycelium (Fig. 6.2)) of up to five spores from each isolate were gently placed over the medium in the following combinations: (DAOM-197198/DAOM-234328; DAOM-197198/DAOM-240415; DAOM-234328/DAOM-240415). Clusters belonging to different isolates were separated by a maximum distance of 2 mm. Plates were then sealed and incubated at 25°C in the dark until germination. Following cluster germination, the occurrence of hyphal fusion was checked weekly for a 45-d period, and results were grouped into four categories (see above, 'Experimental set up' section) (Croll and Sanders 2009). Each experimental unit (i.e. one replicate) consisted of one plate with four repetitions of the same pairing (combination) and ten replicates were performed for each combination. Ten replicates constituted from pairings of the same isolate were used as controls.

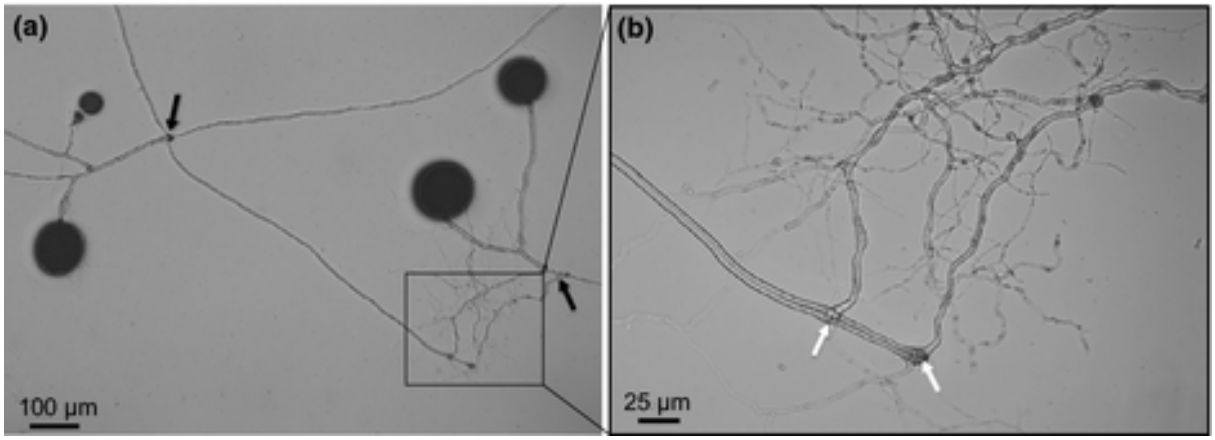


Figure 6.2. Self-fusion between spore clusters belonging to Rhizophagus irregularis DAOM-4240415.
(a) Emergence of several germ tubes through the cut extremities (black arrow) and anastomosis formation. (b) Inset showing details of the anastomosis (white arrows); note the profuse growth of thin hyphae at the interaction zone. Active bidirectional protoplasmic flow, is shown in Video S1.

6.4.7 Symbiotic experiment

An autoclaved (121°C for 15 min) slide (25 9 75 9 1 mm) was placed in the middle of an empty mono-compartment Petri dish (100 9 15 mm). Thirty millilitres of the MM medium was then poured in to fill the plate and cover the slide. Following jellification, only the medium covering the slide was gently removed in such a way that c. 1 mm of each of the slide edges was still embedded in the medium. A flat surface where the development of symbiotic mycelia could easily be traced, adjoined by two identical distal and proximal compartments, was created (Fig. 6.3). Subsequently, 2-cm long Ri T-DNA transformed chicory roots were placed in both compartments and 100 spores of each isolate taken up from monoxenic cultures by solubilisation were inoculated in one of the two compartments using the same combinations described above. Plates were sealed and incubated upside down at 25°C in the dark. Plates were checked weekly and chicory roots growing towards the slide were cut and gently removed using scalpel and forceps. Because mycelia growing on both edges of the slide began to spread across it, plates were checked weekly over 10 wk for hyphal interactions, and results were grouped into four categories according to Croll et al. (2009). Each experimental unit (i.e. one replicate) consisted of one plate inoculated with two isolates in association with a Ri T-DNA transformed chicory root. Fourteen replicates were performed for each combination, and five replicates for each isolate

inoculated in both compartments with the same fungal isolate were used as control.

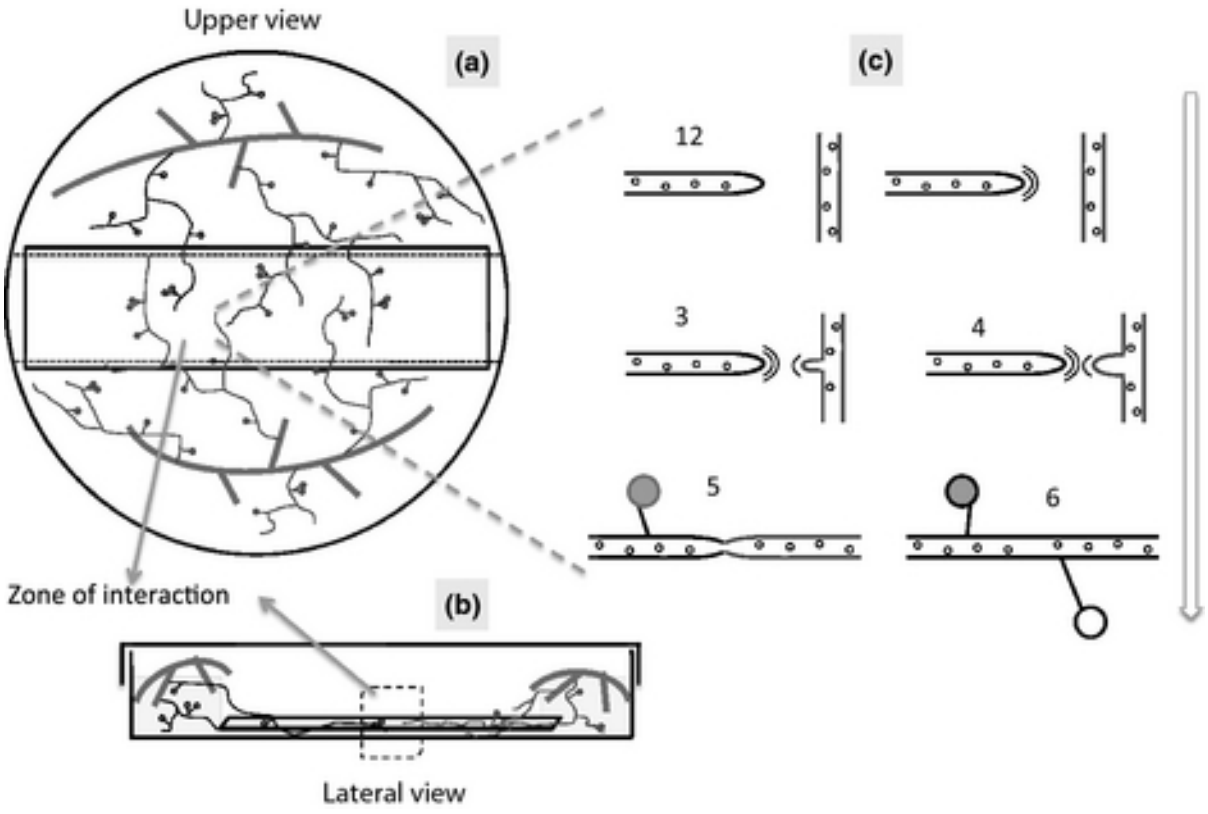


Figure 6.3. Diagram presentation of the experimental set up for the study of anastomosis between geographically distinct isolates of *Rhizophagus irregularis* in the symbiotic mycelium.
(a, b) Different views of the experimental set up showing mycelium (black and gray) from both sides growing towards the slide (i.e. zone of interaction) and (c) interacting with each other.

Monosporal cultures lines with progeny from crossed experiments. In order to test the hypothesis of whether heteroplasmy is a transient state or is maintained into the progeny, twenty full lipid spores (i.e. progenies) harvested from the interaction zone for each combination were

randomly chosen and individually cut out from the mycelium and extracted in 2 µl sterile ultra-pure water (PureLab, Elga, Canada). Spores were placed in a new mono-compartment Petri dish (90 mm) containing M medium in the close vicinity of a Ri-T transformed chicory root. Sixty replicates, each consisting of one single spore associated with a chicory root were maintained as above. Each plate was checked weekly for germination, root colonization and colony development over 11 wk.

6.4.8 Microscopy, data collection, harvest and statistical analysis

For all experiments, the occurrence of anastomosis and/or spore germination was observed under a Discovery V12 stereomicroscope (Carl Zeiss, Toronto, ON, Canada) at magnifications of 96.7–940. Bright-field microscopy (Axio Imager M1; Carl Zeiss) was also used to observe details of hyphal interactions at higher magnifications. Images and movies of the hyphal interactions, as well as protoplasmic streaming through hyphal fusions, were captured with an AxioCam HRC using Axiovision software v4.8.1.0 (Carl Zeiss).

For the second experiment, only the cultures showing synchronized growth of crossed isolates from both sides towards the slide were considered in order to trace the origin of hyphae and to avoid self-fusions. This restriction reduced the experimental replicates to three for some crossing combinations. The weekly growth of the mycelia over the slide was marked with coloured dots according to the isolate, making it easy to trace development as well as interactions between both mycelia. Hyphal interactions, number and phenotype of spores (e.g. shape, lipid droplet content, aborted-like structures) formed at the zone of interaction (slide) at both sides of the fungal cultures were evaluated in addition to measuring total hyphal length and hyphal density. Thirty full lipid spores (20 spores out of 30 were used to initiate monosporal cultures

lines; see experiment 3) formed at the zone of interaction for each combination were randomly chosen, individually cut out from the mycelium and extracted in 2 µl sterile ultra-pure water (PureLab). Whole genome amplification was performed for all spores using the GenomiPhi V2 DNA Amplification Kit (GE Healthcare, Mississauga, ON, Canada) following the manufacturer's instructions. Monosporal culture lines initiated from progenies did not produce any further visible extraradical structures, even though germination and root contact of several spores was observed. Under these circumstances, all germinated and nongerminated spores were harvested and their genome amplified following the same protocol described. The root segment in contact with the germinative mycelium was removed and stained following the protocol described in (Phillips and Hayman 1970) to observe hyphal attachment and also assess the development of intraradical structures. Data analysis was performed using STATISTICA software. Data were subjected to an ANOVA. The Tukey honestly significant difference test was used to identify significant differences ($P \geq 0.05$).

6.4.9 PCR genotyping and sequencing of progeny spores

As already described, isolate-specific markers were designed for each of the three *R. irregularis* isolates, DAOM-240415, DAOM-234328 and DAOM-197198. Primer sequences are summarized in Table 6.1. DNA from individual spore progeny for the second and third experiments was subjected to PCR using the isolate-specific primers corresponding to their parents. PCR conditions were as follows: PCR reactions in a 50-µl volume contained 0.2 mM dNTPs (Fermentas, Canada), 0.5 µM of each primer, 0.5 U of Taq DNA polymerase (BioBasic, Montreal, Quebec, Canada) and 2 µl of the whole genome amplification product of each spore.

PCRs were run on Mastercycler pro S (Eppendorf, Canada). PCR cycling conditions included an initial denaturation at 95°C for 2 min followed by 30 cycles at 95°C for 30 s; 55°C or 60°C for 30 s and 72°C for 30 s. An elongation at 72°C for 10 min was performed at the end of PCR. Hybridization temperature was 55°C for DAOM-240415 and 60°C for DAOM-97198 and DAOM-234328. PCR products were revealed using GelRed dye (Invitrogen) and images were recorded by GelDoc system (BioRad, Canada). PCR was performed according to the crossing combinations summarized in Table 6.1. All the amplification products obtained from progeny were sent for sequencing to McGill University and Genome Quebec Innovation Center facility (Montreal, Canada). Nucleotide Blast searches were performed using the NCBI website to check the presence of the target sequence.

6.5 Results

6.5.1 Mitochondrial genome comparison and marker development

The mtDNAs of the two newly sequenced *R. irregularis* isolates (DAOM-240415 and DAOM-234328) are almost identical, both to each other and to those retrieved from GenBank, *R. irregularis* isolates: 494 (Lee and Young 2009), DAOM-197198 (Nadimi et al. 2012), MUCL-46239, MUCL-46240 and MUCL-46241 (Formey et al. 2012). *Rhizophagus irregularis* isolates have a mitochondrial genome ranging in size from 68 995 bp to 74 799 bp, except for the isolate MUCL-43204, which is outside of this range (87 754 bp). This variation in size is caused by a divergence in the presence/absence and size of mobile elements such as plasmid related DNA polymerase genes (*dpo*), open reading frames encoding homing endonuclease genes in defined intergenic regions, as previously described in Formey et al. (2012).

In order to develop the mtDNA isolate-specific marker for DAOM-197198, DAOM-240415 and DAOM-234328, multiple alignments of their complete mtDNAs and those of five other *R. irregularis* isolates were performed. We focused on the *cox3-rnl* intergene, which is a large region comprising substantial sequence divergence, as a target to design size-variable isolate-specific markers. The specific markers that were developed were used to distinguish between the three *R. irregularis* isolates selected in our crossing experiments, although the presence of a large number of *dpo* elements of different lengths and orientations in the *cox3-rnl* intergenic region made it challenging to perform the multiple sequence alignments. Each specific marker amplified a single band of an expected size for the corresponding isolate (Fig. 6.1). In addition to the specificity, the three markers were designed to produce different PCR product sizes, in order to facilitate tracing each marker in spore progeny. For example, marker A, specific for isolate DAOM-197198, gave a PCR band of 120 bp only for DNA of its corresponding isolate and no band was observed for isolates DAOM-240415 and DAOM-234328. Markers B and C, specific for isolates DAOM-240415 and DAOM-234328, gave PCR bands of 147 bp and 209 bp, respectively (Fig. 6.1, Table 6.1). Length polymorphism can also be used to trace each haplotype in multiplex PCR reactions. However, primers of isolate DAOM-234328 can also amplify mtDNA of the isolate 494, producing a 209-bp PCR product. Because isolate 494 was not included in our crossing experiment, the primer pair developed for DAOM-234328 was valid.

6.5.2 Pre-symbiotic interactions

In this experiment, we recorded both self and nonself fusion between all the isolates tested and characterized them by total fusion of hyphal membranes and cell wall followed by bidirectional protoplasm streaming between the connected hyphae (Video S6.1). Cluster

germination for all isolates was observed during the first 24 h. They germinated as a syncytium, which means that the germ tube never emerged through the subtending hypha when the spores were sharing a common mycelium, but one or several germ tubes re-grew from the cut extremities (Fig. 6.2). After 1 wk, several spores within the cluster lost their cytoplasmic content and became translucent, while others within the same cluster seemed to be dormant, as no change in color and lipid content was observed.

Self-fusion was a common event within the same isolate (Fig. 6.2), and ranged from 60% to 38% in DAOM-197198 and DAOM-2404515, respectively. Pre-fusion as well as post-fusion incompatibility was never recorded in pairings within the same isolate. On the contrary, hyphal interaction between pairings of different isolates (i.e. nonself interaction) showed heterogeneous responses. Nonself hyphal fusion was rare between different isolates, and was recorded only two times in the combination DAOM-234328/DAOM-240415 out of 157 contacts observed (Supporting Information Table S6.1). Nonself post-fusion incompatibility was recorded for all combinations and the highest value was 3 (2%) out of 148 contacts observed in DAOM-197198/ DAOM-240415. Microscopic observations of these hyphal fusions revealed slow protoplasm movement alongside the connected hyphae followed by protoplasm withdrawal and septa formation (Video S6.2).

6.5.3 Symbiotic interactions

Growth of the mycelium during the symbiotic phase showed the same developmental patterns for *Rhizophagus* species (formerly *Glomus* genus) (Bago et al. 1998). After germination and root colonization, runner hyphae (RH) rapidly colonized the medium, and extended towards the interaction zone. They divided dichotomously into thinner hyphae of increasing branch order. Hyphae subsequently branched profusely to form the so-called branching absorbing structures

(BAS). Analysis of the hyphal interactions between isolates also showed a heterogeneous response, as in the first experiment. No differences in hyphal density, number of contacts or anastomosis formation between the treatments were detected ($P \geq 0.05$) (Table 6.2). Perfect hyphal fusions between colonies of the same isolate ranged from 19 ± 8.2 in DAOM-240415 to 49.5 ± 17.5 in DAOM-234328 (mean ± SE) Pre-fusion as well as post-fusion incompatibility was never recorded, as observed in the pre-symbiotic interactions.

Table 6.2. Anastomosis[1,2] frequency from the interaction of *Rhizophagus irregularis* isolates in the symbiotic experiment

Type of interaction	Hyphal density at the interaction zone[3] (cm cm^{-2})	Number of contacts	Perfect fusion (self fusion)[4]	No interaction	Pre-fusion incompatibility or hyphal avoidance	Post-fusion incompatibility
Within the same isolate						
A–A	10.8 ± 1	42 ± 25	26 ± 23	16 ± 2	0	0
B–B	13.4 ± 3.1	55 ± 14	19 ± 8.2	36.3 ± 6.6	0	0
C–C	14.6 ± 1.2	77 ± 18	49.5 ± 17.5	30 ± 3	0	0
Between isolates	**Hyphal density at the interaction zone (cm cm^{-2})**	**Number of contacts**	**Nonself fusion[d]**	**No interaction**	**Nonself pre-fusion incompatibility**	**Nonself post-fusion incompatibility**
A–C	9.4 ± 1.4	25 ± 7.1	0	23.7 ± 7.8	1.2 ± 0.8	0
A–B	12 ± 2	25 ± 7.1	0	23.8 ± 7.7	1.25 ± 0.7	0
B–C	11.2 ± 1.6	26.7 ± 4.3	1.3 ± 0.9	25.3 ± 3.9	0.33 ± 0.33	0

A, DAOM 197198; B, DAOM 240415; C, DAOM 234328.

[1] Anastomoses within the same mycelium in crossing treatment were not recorded.

[2] No significance differences were found within the treatments and controls ($P \geq 0.05$).

[3] Mean ± SE.

[4] Fusions were followed for 3 wk and no septa were formed between the connected hyphae.

Nonself fusions were only observed at the zone of interaction for the combination DAOM-240415/DAOM-234328. Because anastomosis traceability was solely confined to the zone of the interaction, the occurrence of nonself fusions far beyond the interaction zone might not be ruled out. Pre-fusion incompatibility was observed for all combinations characterized by hyphal burst at the hyphal apex or septa formation and cytoplasm retraction before contact. No signs of wall thickening at the hyphal-apex region were observed when hyphae from different isolates came into contact, but overlapped mycelium was the rule rather than the exception for all combinations. Sporogenesis was observed during the first week after root colonization on the RH. Most spores were formed into lower order hyphae after extensive growth of the fungal colony. Full lipids and empty spores as well as aborted-like structures were observed and recorded in all the isolates tested, regardless of the combination. The percentage of aborted-like structures in the interaction zone ranged from absence to 2 ± 1(mean ± SE) among crossed experiments and from 2 ± 2 to 5 ± 2 in controls. The same pattern was observed at each side, when the plate was inoculated with both isolates in crossing (i.e. nonself interaction), or within controls (i.e. self interactions). Statistical analysis did not show any significant difference in the number and percentage of spores produced for any of the categories ($P \geq 0.05$) mentioned above (Table S6.2). Although in the combination DAOM-197198/DAOM-240415 (Fig. 6.4, Video S6.3), several aborted-like structures (data not shown) often arrested their development and burst, after which several hyphae grew out of the spore lumen (Fig. 6.5) Observations of these hyphae showed that cytoplasm retracted and several septa were then formed.

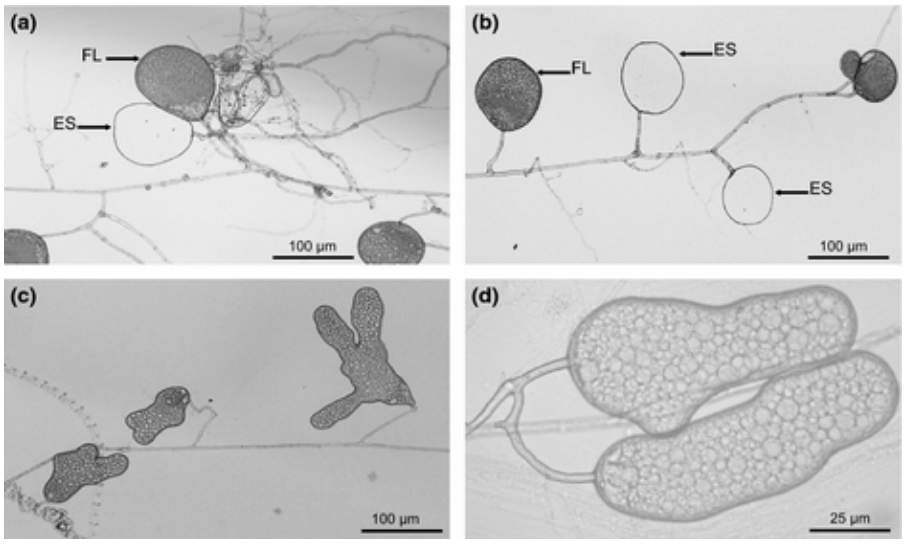


Figure 6.4. Variability in spore morphology of *Rhizophagus irregularis* at the interaction zone in the combination DAOM197198/DAOM240415.

(a) Spore clusters showing empty, full lipid droplets and crushed spores. (b) Extraradical mycelium bearing both full lipid droplets (FL) and emptied spores (ES). (c, d), Irregular shaped spores within the same mycelium. For a detailed view of the spore variability see Video S3.

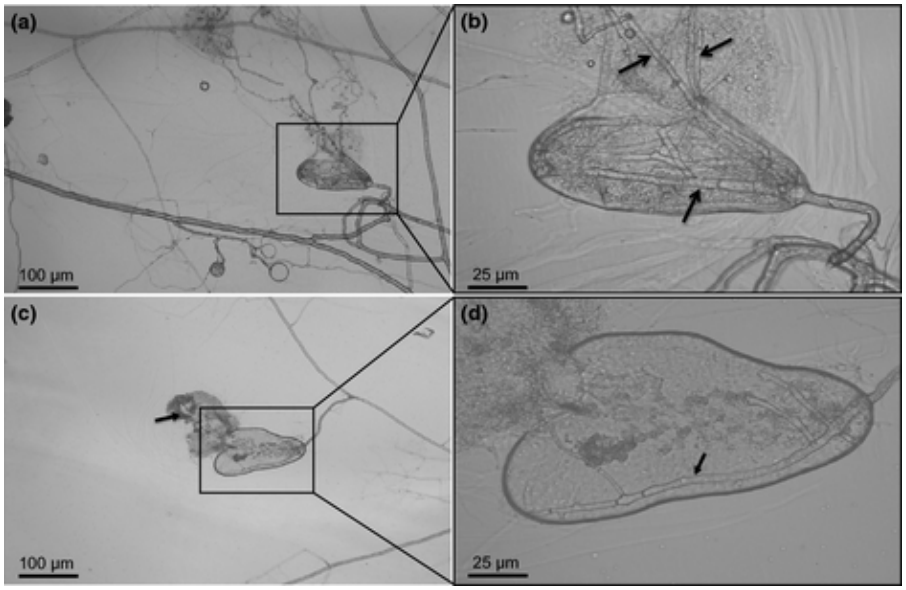


Figure 6.5. Aborted-like structures produced by the *Rhizophagus irregularis* isolates at the interaction zone.
(a, b) DAOM-197198/DAOM-240415; (c, d) DAOM-234328/DAOM-240415. (a) Several hyphae exit from the aborted-like structures lumen and continue growing following cytoplasm retraction and multiple septa formation. (b) Inset showing details of the lumen and septated hyphae (arrows). (c) Burst and cytoplasm exit (black arrow). (d) Inset showing details of the septated hypha inside the aborted-like structures.

6.5.4 mtDNA genotyping and sequencing in spore progeny

For each crossing combination in a symbiotic interaction, 10 full-lipid spores of the progeny were individually genotyped using isolate-specific markers corresponding to their parents, as well as the progeny of the monosporal cultures lines experiment. The progeny of the three crossing combinations showed different PCR amplification patterns (Fig. 6.6). For example, the crossing of isolates DAOM-197198 and DAOM-240415 showed that nine progeny spores (S) out of ten contained both parental markers and one spore (S6) did not produce a PCR amplification product using mtDNA markers, while 18S rDNA primers AML1 and AML2 showed a PCR band of an expected size (790 bp; data not shown). Five spores of the crossing combination between DAOM-197198 and DAOM-234328 showed both mtDNA haplotypes, while five contained only 197198-mtDNA haplotype. Three spores (S3, S6 and S9) of the crossing combination between DAOM-240415 and DAOM-234328 showed both mtDNA haplotypes, six spores showed one mtDNA haplotype whereas DAOM-240415 mtDNA haplotype was predominant (S1, S2, S5, S7 and S10) and spore (S8) did not produce a PCR amplification product using mtDNA markers, while 18S rDNA primers did produce a PCR band (data not shown). Interestingly, the PCR bands varied greatly in intensity among spores and independently of the length of the marker, probably due to the frequency of each mtDNA haplotype.

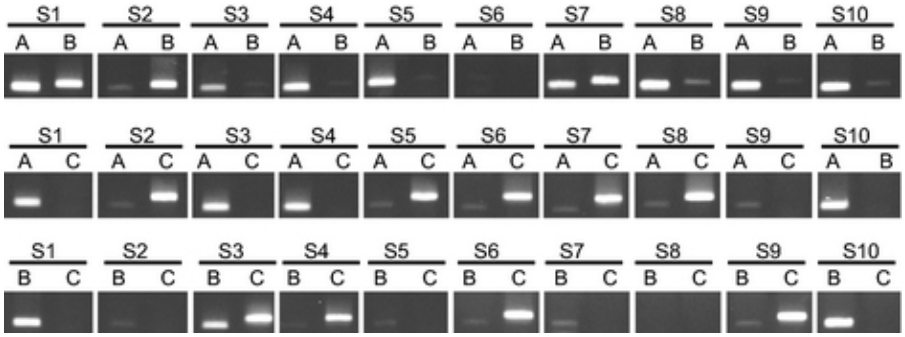


Figure 6.6. Gel electrophoresis showing patterns of mtDNA genotyping of ten progeny spores (S1–S10) of three combinations using three *Rhizophagus irregularis* isolates DAOM-197198 (A) DAOM-240415 (B) and DAOM-234328 (C).
mtDNA marker sizes are 120, 147 and 209 bp corresponding to DAOM-197198, DAOM-240415 and DAOM-234328, respectively.

6.5.5 Monosporal cultures with progeny from crossed experiments

Even though attempts to obtain monosporal culture lines failed, spore germination was observed for progenies from all crossed experiments and ranged from 40% in DAOM-197198/DAOM-234328 to 60% in DAOM-240415/DAOM-234328 (Table S6.3). The germination process was characterized by the emission of one to several germinative hyphae through the subtending hyphae (Fig. S6.1). Some germinative hyphae showed a marked rhizo-tropism (Fig. S6.1); however, a nonorientated growth was also observed. Germinative hyphae also underwent contact with roots but after several weeks no progress of the fungal growth was observed (i.e. runner hyphae bearing spores and branching absorbing structures).

Genotyping of the progeny showed that 14 spores gave positive PCR amplifications while the rest of the progeny did not produce any amplification product. All PCR products were

sequenced and matched to the isolate-specific parental mtDNA markers. However, genotyping results showed that all the progeny amplified only one parental mtDNA haplotype (Fig. S6.2). For example, in the combination DAOM-197198/DAOM- 234328, three spores were genotyped of which two contained the DAOM-197198-mtDNA haplotype and the other showed the DAOM-234328 mtDNA haplotype. Five spores in the combination DAOM-240415/DAOM-234328 were genotyped where two spores contained the DAOM-234328 mtDNA and three presented the DAOM-240415 mtDNA haplotype. Finally, six spores were genotyped in combination DAOM-197198/ DAOM-240415, where only one showed the DAOM-197198 genotype while the others contained the DAOM-240415 mtDNA haplotype.

6.5.6 Colonization detection

Microscopic observation of stained root fragments, which were suspected for being colonized by progenies, revealed the absence of hyphopodia formation. Consequently, no intraradical structure was observed.

6.6 Discussion

Nonself recognition between members of the *Rhizophagus* genus has been documented (Croll and Sanders 2009;Angelard et al. 2010; Colard et al. 2011; Purin and Morton 2012) and has resulted in nuclear segregation to new individuals affecting plant fitness (Angelard et al. 2010; Colard et al. 2011). Interestingly, combination B-C (isolates DAOM-240415 originating from Manitoba, Canada, and DAOM-234328 originating from Finland) showed perfect nonself fusion and the occurrence of heteroplasmy. Moreover, in the symbiotic crossing experiment of

combinations A-B and A-C, we did not observe perfect nonself fusion, but a few proportions indicative of pre-fusion incompatibility; genotyping of mtDNA clearly showed the presence of both parental haplotypes in the spore progeny. These results can likely be explained on the basis of the coenocytic nature of the AMF mycelium allowing protoplasm mixture before hyphal septation, suggesting the existence, as in higher fungi (Fu et al. 2011), of a putative molecular mechanism regulating nonself recognition that has yet to be demonstrated.

The observed mtDNA heteroplasmy in the progeny following nonself recognition (i.e. unrelated to apparent sexual reproduction) between genetically divergent isolates of the same species originating from different continents is somewhat surprising. This assumption is made on the basis that all mtDNA sequenced so far in AMF (Borstler et al. 2008; Lang and Hijri 2009; Lee and Young 2009; Formey et al. 2012; Nadimi et al. 2012; Pelin et al. 2012; Beaudet et al. 2013) and those presented in this study, have been shown to be homogeneous. This apparent contradiction is evidence of the putative transient heteroplasmic state in AMF, and is consistent with our results as both germinated progeny and spores that do not break their dormancy after long periods of sub-cultivation showed the presence of one parental mtDNA marker. This evidence suggests the presence of an effective segregation mechanism, which might regulate mtDNA inheritance in these organisms. Inheritance of mitochondria in fungi during the mating process is mostly uniparental, with only one parental haplotype of the mitochondrial genome inherited by the progeny (reviewed in (Ni et al. 2011)). However, biparental inheritance of mtDNA has also been shown in some fungal species, such as *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, because the two mating cells contribute equally through their cytoplasm organelles to form a mixed cytoplasm in the zygote. Interestingly, the coexistence of different mitochondrial genome haplotypes has been shown to be transient in these organisms (Birky 2001; Ni et al. 2011; Beaudet et al. 2013). Mechanisms contributing to homoplasmy are

often related to the recombination of both parental mtDNAs or by the selection of one of the parental haplotypes due to the presence of segregation mechanisms, such as genetic bottleneck followed by a stochastic drift (Birky 2001). As proposed by White et al. (2008), heteroplasmy may either persist, become fixed, or be lost due to either selection forces or drift, or both. The extent to which such mechanisms may occur in AMF and the resulting haplotype to be fixed in the population will need to be assessed.

In arbuscular mycorrhizal fungi, no studies have been addressed to investigate patterns of mtDNA inheritance and segregation to the progeny, however, the occurrence of distinct mtDNA haplotypes within an AMF individual, with divergence in mobile elements and *dpo* insertions, points to the possible occurrence of interhaplotype recombination and mtDNA segregation mechanisms in these fungi (D. Beaudet et al. unpublished data). In budding yeasts mtDNA segregation occurs via nucleoids (i.e. packages of identical mtDNAs maintained together by Holliday junctions and proteins; (Lockshon et al. 1995; White and Lilley 1997; MacAlpine et al. 1998), creating a genetic bottleneck responsible for further rapid segregation. Due to this mechanism, heteroplasmy can last over 20 generations in these organisms. It would be interesting to investigate whether the formation of such mitochondrial nucleoids occurs in AMF and, if so, how long this heteroplasmic state is maintained. The efficiency of the mitochondrial segregation process is directly correlated to the initial population size and the relative proportion of one given haplotype in a population (White et al. 2008). In this study, only two different mtDNA haplotypes were mixed *in vitro*, giving rise to limited initial genetic variants. Also, we only have evidence of the mtDNA haplotypes frequency from the intensity of the PCR bands we observed, which make it difficult to predict the effectiveness of the segregation process.

Other studies have shown that coexistence of two mtDNAs within a common cytoplasm may be disadvantageous, likely due to genetic conflicts or competition between divergent mtDNA

haplotypes (reviewed in (Ni et al. 2011), references therein). It has been proposed that uniparental inheritance of mtDNA in fungi as well as other eukaryotes might be maintained by selection forces to avoid such conflicts (White et al. 2008). Some authors have hypothesized that uniparental inheritance of mtDNA and chloroplastic DNA may avoid dispersion of selfish and deleterious mobile elements (Hoekstra 2000; Mamirova et al. 2007). Even if the divergence between the mtDNA haplotypes used in this study appeared to be neutral, because they are located in intergenic regions, it is difficult to predict with confidence the implied function in the genome and its role in the segregation and inheritance process. In any case, selection of one mtDNA haplotype due to potential mtDNA conflicts or stochastic segregation mechanisms might explain the transient character of heteroplasmy in our experiment and also provides a rational explanation for why homoplasmy is more often observed in spores from field (Borstler et al. 2008) and/or *in vitro* experiments than heteroplasmy (Lee and Young 2009; Formey et al. 2012; Nadimi et al. 2012; Pelin et al. 2012; Beaudet et al. 2013). Previous studies on AMF mtDNA have already demonstrated the great potential for developing isolate-specific molecular markers (Borstler et al. 2008; Lang and Hijri 2009; Formey et al. 2012; Beaudet et al. 2013). For example, Borstler et al. (2008) used the *rnl* gene encoding the mitochondrial ribosomal large subunit (mtLSU) to discriminate different haplotypes among cultivated isolates of *R. irregularis* (formerly *G. intraradices*) in soil and within root samples from the field. These authors found 12 *rnl* haplotypes among 16 isolates of *R. irregularis* originating from five continents. Formey et al. (2012) recently compared the complete mtDNA of six *R. irregularis* isolates and were able to develop mtDNA markers using the *cox3-rnl* intergene as well as other regions to discriminate isolates using length polymorphism markers. These mtDNA markers can be useful for other studies, both in population genetics as well as in tracing isolates in agriculture soil and plant roots to test co-inoculation of different *R. irregularis* isolates. However, it will be important to have a

better understanding of the mitochondrial inheritance process and the extent of the mtDNA variability present in a natural population in order to benefit from the full potential of such markers.

In this particular study, we believe it was crucial to use size-variable markers combined with a whole genome amplification step, to easily test the occurrence of heteroplasmy between the crossed parental cultures. Some studies have suggested that during their formation, spores can randomly inherit different nuclear haplotypes following hyphal fusion (Marleau et al. 2011). The same pattern could be expected for mtDNA, however they could be critically underrepresented in daughter spores, given the competition with the endogenous haplotypes and the putative process of mtDNA segregation. If the isolate-specific markers were variable in single nucleotide polymorphisms (SNPs), it would have required a large number of clones for each tested spore to find even one different locus. PCR and cloning-based studies with such markers could fail to identify occurrences of heteroplasmy. Despite its potential as a DNA barcode, mtDNA has never been considered as a unique and reliable highly-sensitive target to address studies on fungal individualism in AMF (Rayner 1991).

6.7 Conclusions

We have demonstrated a potential use of mtDNA markers to address basic questions of evolutionary biology in AMF using *R. irregularis* isolates as a model. Our study provided additional information in the AMF mitochondrial genome collection with the complete mtDNA of two *R. irregularis* isolates. Currently, there are eight *R. irregularis* isolates with annotated mitochondrial genomes that differ substantially in some intergenic regions, making them useful sites for developing isolate-specific markers. We also provided additional markers that were

specific to some isolates, and demonstrated their potential use in tracing mtDNA in a crossing experiment.

We also showed that isolates of *R. irregularis* from different geographical locations, even from different continents, were able to exchange their genetic material regarding mtDNA, resulting in heteroplasmy in their spore progeny. The extent to which such a state may occur in natural communities and the mechanisms responsible for the fixation of a homogeneous haplotype in an AMF individual need to be assessed. Further studies on vegetative compatibility and incompatibility as well as putative sex machinery in AMF, are needed to advance our understanding of the evolution of AMF and to better understand their population dynamics.

6.8 Acknowledgements

This study was conducted as part of a research project organized and coordinated by Premier Tech. The authors are grateful for financial supports from NSERC Cooperative Research and Development (grant number: RDCPJ 395241-09), NSERC Discovery (grant number: 328098-2012), Premier Tech and CRI-BIQ. I.D.P. was supported by Genome Canada and Genome Quebec funds. We would also like to thank Biopterre centre du developpement des bioproduits and CRBM for assistance and Nathalie De Jaeger, Terrence Bell, Karen Fisher Favret and three anonymous referees for their useful reading and comments on the manuscript.

Chapter 7 - General discussion, Conclusion and Perspectives

My Ph.D. thesis contributes significantly to fundamental knowledge on the evolution of mitochondrial genomes in fungi, in particular in the early diverging fungal lineage of Glomeromycota. It clearly reports the presence of substantial intra- and inter-specific genetic variation in glomeromycotan mtDNA. My thesis shows that these substantial variations are generally caused by the invasion and integration of mobile elements such as plasmid-related DNA polymerase gene (*dpos*), open reading frames (*orfs*) and small inverted repeats (SIRs). I showed that mobilome variation and movement within mt DNA of AMF contributed to the evolution of unorthodox genes, genome structure and variable intergenic and intronic regions. The outcome of mobilome variation in mt genomes of AMF (e.g. genome rearrangement) can be exploited as opportunities for developing efficient and accurate isolate- and species-specific molecular marker. For example, formation of fragmented genes in *Gigaspora rosea* (chapter 2), unorthodox C-terminal regions in some mt protein-coding genes (PCGs) of some *Rhizophagus* spp. isolates (Chapter 3) and the mt genome fragmentation in an isolate of *Rhizophagus* sp. (DAOM 213198) (Chapter 5), are mediated by mobile element activities. Consequently, I hypothesize that the biochemical machineries for both homologous and non-homologous recombination might be active in mitochondria of glomeromycotan fungi, although, this has not been demonstrated yet.

The reported evolutionary mechanisms that are mediated by mobilome activities (*dpos*, *orfs* and SIRs) also increase genetic polymorphism in intergenic regions giving rise to hypervariable

regions or “hotspots” for example *cox3-rnl* intergene. This intergene varies substantially among closely related species and even among different isolates belonging to the same species as is shown for the model AMF species *R. irregularis*, although conserved intergenic regions exist. The genetic polymorphism observed in mtDNAs has been shown to be useful to isolate-specific biomarkers development, as it has been applied and presented in chapters 3 and 4. Accurate biomarkers and diagnosis toolkits are useful for identification, characterization and quantification not only in AMF but also in other organisms. These studies have potential fundamental implications in studying population genetics, ecology and functions of Glomeromycetes in ecosystems, but also in applied research in agriculture and environmental protection.

In my Ph.D. study, I designed, tested and validated isolate-specific molecular markers for the study of genetic exchange in three geographically distant isolates of *R. irregularis*. Using these isolate-specific markers, I identified their efficiency in tracing and tracking mitochondria within fused AMF isolates through anastomosis (chapter 5). My markers allowed discovering a transient heteroplasmy in the AMF *R. irregularis*. I showed that isolates of *R. irregularis* from different geographical locations can exchange their genetic material regarding mtDNA, resulting in heteroplasmy in their spore progenies. However, this heteroplasmy state appeared to be transient because of its loss in progeny spores that were able to germinate. This evidence suggests the presence of an effective segregation mechanism, which might regulate mtDNA inheritance in AMF and so support the reliability of mitogenome utilization for biomarker development. However, since the divergence between the mtDNA haplotypes used in our study appeared to be neutral (i.e. divergence mediated by *dpo* within hypervariable region of *cox3-rnl*), it is difficult to predict the implied function in the genome and its role in the segregation and inheritance process. Further experiments are needed to address anastomosis among isolates harboring mtDNA haplotypes vary in the coding sequence. Such length specific markers and isolate-specific

markers could also be used to investigate population genetics and interaction of AMF with plant roots and soil microorganisms.

In my Ph.D. thesis, I contribute significantly to the implementation of completed mtDNAs in public databases and comparing mt genomes among different AMF taxa. My study provides additional information in the AMF mitochondrial genome collection with the complete mtDNA of ten *R. irregularis* isolates. I also performed phylogenetic analyses to compare and assess phylogenetic signals between single mt genes, subsets of genes and complete set of concatenated mt genes (16 mt coding genes). Additionally I documented an unusual mtDNA organization in a close relative to the model species *R. irregularis* and this finding will lead to a new level of understanding of mt genome evolution in eukaryotes. Moreover, I report that from a phylogenetic perspective, not all mt genes possess the same phylogenetic signal and thus can not be equally representative of the complete mt genes set.

Our phylogenetic analyses of mt protein-coding genes revealed that AMF are related to Mucoromycotina rather than to the Dikarya phyla (Basidiomycota and Ascomycota), which was in contrast to rDNA analyses that was suggesting the Zygomycota phylum is paraphyletic and that Glomeromycota are sister to the Dikarya phyla (Schüßler et al. 2001; James et al. 2006). However, recent findings revealed that the 45S rDNA locus of *R. irregularis* is highly polymorphic (Stockinger et al. 2009; Lin et al. 2014) and so is not a proper candidate for such analyses. Further analyses of a supermatrix of 35 highly conserved, putative single copy nuclear genes (Capella-Gutie´rez et al. 2012) support our result. Consequently, similar results in revealing evolutionary relationship of AMF within Fungi support our hypothesis that mtDNA can be a proper compliment for identification and phylogenetic analyses comprising relatively easier to access database.

Sequencing of other mtDNAs in distant lineages of AMF in order to cover major AMF taxa (families and genera) will advance our knowledge of this important fungal group. It will also open new opportunities for phylogenetic and identification analyses in a wider taxonomic level, which will help to improve AMF function in agriculture and forestry. However, sequencing mtDNAs may be challenging in some AMF taxa, in particular, those that are not able to grow in vitro, or the ones with low sporulation rate and slow growth. To overcome these limitations development or optimization of cultivation both in vitro and in vivo conditions, could improve sequencing procedure and bioinformatic analyses. Further studies need to reevaluate mt genes phylogenetic signal at different taxonomic levels and within distant lineages of AMF. Taken all together, comparison and complementing of mitochondrial- and nuclear-based phylogenetic analyses and identification will help us better understand and apply AMF in nature. Moreover, advanced investigation of mobilome (e.g., origin of mobilome, their replication and dispersal mechanisms, their affinity to integrate specific regions such as hot-spots) that resides within mt genomes of AMF will decipher their contributions in AMF mtDNA evolution.

Broad taxa sampling for mt genomic and phylogenetic analyses will strengthen anastomosis and mtDNA exchange studies among AMF. This could lead to discovering mechanisms and machineries that control vegetative compatibility and incompatibility. Implementation of mtDNA datasets with morphological characters using advanced microscopy techniques and staining procedures will aid taxonomists to better classify and study the evolutionary relationships in Glomeromycata.

In conclusion, my thesis, along with other studies on mtDNA, report the usefulness of mitochondrial genomics to address basic and applied research questions in glomeromycotan biology. Accurate mt biomarkers and diagnostic toolkits have proven their efficiency and could be used in applied aspects, such as the identification and quantification of AMF, their in vitro or

in vivo cultivation (e.g., in bioremediation applications or in agriculture investigations). These tools could ultimately, improve the efficiency of phytoremediation and also sustain agriculture by better managing AMF-based biofertilizers.

References

Adams, K. L., M. J. Clements and J. C. Vaughn (1998). "The Peperomia mitochondrial coxI group I intron: timing of horizontal transfer and subsequent evolution of the intron." J Mol Evol **46**(6): 689-696.

Aguilera, P., T. Barry and J. Tovar (2008). "Entamoeba histolytica mitosomes: organelles in search of a function." Exp Parasitol **118**(1): 10-16.

Aliasgharzad, N., M. R. Neyshabouri and G. Salimi (2006). "Effects of arbuscular mycorrhizal fungi and Bradyrhizobium japonicum on drought stress of soybean." Biologia **61**: S324-S328.

Alkan, N., V. Gadkar, O. Yarden and Y. Kapulnik (2006). "Analysis of quantitative interactions between two species of arbuscular mycorrhizal fungi, Glomus mosseae and G. intraradices, by real-time PCR." Appl Environ Microbiol **72**(6): 4192-4199.

Alverson, A. J., D. W. Rice, S. Dickinson, K. Barry and J. D. Palmer (2011). "Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber." Plant Cell **23**(7): 2499-2513.

Andronescu, M., A. Condon, H. H. Hoos, D. H. Mathews and K. P. Murphy (2007). "Efficient parameter estimation for RNA secondary structure prediction." Bioinformatics **23**(13): i19-28.

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

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

Angelard C, Tanner CJ, Fontanillas P, Niculita-Hirzel H, Masclaux F, Sanders IR. (2014). Rapid genotypic change and plasticity in arbuscular mycorrhizal fungi is caused by a host shift and enhanced by segregation. ISME J 8:284-294.

Anziano, P. Q., D. K. Hanson, H. R. Mahler and P. S. Perlman (1982). "Functional domains in introns: trans-acting and cis-acting regions of intron 4 of the cob gene." Cell **30**(3): 925-932.

Armstrong, M. R., M. S. Blok Vc Fau - Phillips and M. S. Phillips "A multipartite mitochondrial genome in the potato cyst nematode Globodera pallida." (0016-6731 (Print)).

Avise, J. C. (1994). Molecular markers, natural history and evolution. New York, Chapman & Hall.

Azcón-Aguilar, C. and J. M. Barea (1997). "Arbuscular mycorrhizas and biological control of soil-borne plant pathogens – an overview of the mechanisms involved." Mycorrhiza **6**(6): 457-464.

Bago, B., C. Azcón-Aguilar and Y. Piché (1998). "Architecture and Developmental Dynamics of the External Mycelium of the Arbuscular Mycorrhizal Fungus Glomus intraradices Grown under Monoxenic Conditions." Mycologia **90**(1): 52-62.

Baidyaroy, D., J. M. Glynn and H. Bertrand (2000). "Dynamics of asexual transmission of a mitochondrial plasmid in Cryphonectria parasitica." Curr Genet **37**(4): 257-267.

Ballard, J. W. and A. C. James (2004). "Differential fitness of mitochondrial DNA in perturbation cage studies correlates with global abundance and population history in Drosophila simulans." Proc Biol Sci **271**(1544): 1197-1201.

Barea, J. M., R. Azcon and C. Azcon-Aguilar (2002). "Mycorrhizosphere interactions to improve plant fitness and soil quality." Antonie Van Leeuwenhoek **81**(1-4): 343-351.

Barr, C. M., M. Neiman and D. R. Taylor (2005). "Inheritance and recombination of mitochondrial genomes in plants, fungi and animals." New Phytol **168**(1): 39-50.

Beaudet, D., M. Nadimi, B. Iffis and M. Hijri (2013). "Rapid mitochondrial genome evolution through invasion of mobile elements in two closely related species of arbuscular mycorrhizal fungi." PLoS One **8**(4): e60768.

Beaudet, D., Y. Terrat, S. Halary, I. E. de la Providencia and M. Hijri (2013). "Mitochondrial genome rearrangements in glomus species triggered by homologous recombination between distinct mtDNA haplotypes." Genome Biol Evol **5**(9): 1628-1643.
Becard, G. and J. A. Fortin (1988). "Early events of vesicular–arbuscular mycorrhiza formation on Ri T-DNA transformed roots." New Phytologist **108**(2): 211-218.
Bell-Pedersen, D., S. Quirk, J. Clyman and M. Belfort (1990). "Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications." Nucleic Acids Research **18**(13): 3763-3770.
Benard, G. and M. Karbowski (2009). "Mitochondrial fusion and division: Regulation and role in cell viability." Semin Cell Dev Biol **20**(3): 365-374.
Bendich, A. J. "Structural analysis of mitochondrial DNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis." (0022-2836 (Print)).
Bertrand H Fau - Griffiths, A. J., D. A. Griffiths Aj Fau - Court, C. K. Court Da Fau - Cheng and C. K. Cheng "An extrachromosomal plasmid is the etiological precursor of kalDNA insertion sequences in the mitochondrial chromosome of senescent neurospora." (0092-8674 (Print)).
Bever, J. D. and M. Wang (2005). "Arbuscular mycorrhizal fungi: Hyphal fusion and multigenomic structure." Nature **433**(7022): E3-E4.
Bever, J. D. and M. Wang (2005). "Arbuscular mycorrhizal fungi: hyphal fusion and multigenomic structure." Nature **433**(7022): E3-4; discussion E4.
Biessmann, H., K. Valgeirsdottir, A. Lofsky, C. Chin, B. Ginther, R. W. Levis and M. L. Pardue (1992). "HeT-A, a transposable element specifically involved in "healing" broken chromosome ends in Drosophila melanogaster." Molecular and Cellular Biology **12**(9): 3910-3918.
Birky, C. W., Jr. (2001). "The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models." Annu Rev Genet **35**: 125-148.
Blaszkowski J, C. B., Wubet T, Schuessler T, Buscot F, Renker C. (2008). "Glomus irregulare, a new arbuscular mycorrhizal fungus in the Glomeromycota." Mycotaxon **106**: 247-267.
Boer, P. H. and M. W. Gray (1988). "Scrambled ribosomal RNA gene pieces in Chlamydomonas reinhardtii mitochondrial DNA." Cell **55**(3): 399-411.
Bolan, N. S. (1991). "A critical review on the role of mycorrhizal fungi in the uptake of phosphorus by plants." Plant and Soil **134**(2): 189-207.
Bonen, L. (2008). "Cis- and trans-splicing of group II introns in plant mitochondria." Mitochondrion **8**(1): 26-34.
Boon, E., E. Zimmerman, B. F. Lang and M. Hijri (2010). "Intra-isolate genome variation in arbuscular mycorrhizal fungi persists in the transcriptome." Journal of Evolutionary Biology **23**(7): 1519-1527.
Boon, E., E. Zimmerman, M. St-Arnaud and M. Hijri (2013). "Allelic differences among sister spores suggest genetic drift in an arbuscular mycorrhizal fungus." PLoS ONE **8**: e83301.
Borriello, R., V. Bianciotto, A. Orgiazzi, E. Lumini and R. Bergero (2014). "Sequencing and comparison of the mitochondrial COI gene from isolates of Arbuscular Mycorrhizal Fungi belonging to Gigasporaceae and Glomeraceae families." Mol Phylogenet Evol **75**: 1-10.
Borstler, B., P. A. Raab, O. Thiery, J. B. Morton and D. Redecker (2008). "Genetic diversity of the arbuscular mycorrhizal fungus Glomus intraradices as determined by mitochondrial large subunit rRNA gene sequences is considerably higher than previously expected." New Phytol **180**(2): 452-465.
Borstler, B., P. A. Raab, O. Thiery, J. B. Morton and D. Redecker (2008). "Genetic diversity of the arbuscular mycorrhizal fungus Glomus intraradices as determined by mitochondrial large subunit rRNA gene sequences is considerably higher than previously expected." New Phytologist **180**(2): 452-465.
Brown, G. G. and M. Zhang (1995). Mitochondrial plasmids: DNA and RNA. The molecular biology of plant mitochondria. V. I. e. Levings CS III. Dordrecht, Kluwer**:** 61-91.

Brügger, K., E. Torarinsson, P. Redder, L. Chen and R. Garrett (2004). "Shuffling of Sulfolobus genomes by autonomous and non-autonomous mobile elements." Biochemical Society Transactions **32**: 179-183.
Brundrett, M. (2004). "Diversity and classification of mycorrhizal associations." Biological Reviews **79**(3): 473-495.
Bruns, T. D., D. S. Palmer Jd Fau - Shumard, L. I. Shumard Ds Fau - Grossman, M. E. Grossman Li Fau - Hudspeth and M. E. Hudspeth (1988). "Mitochondrial DNAs of Suillus: three fold size change in molecules that share a common gene order." (0172-8083 (Print)).
Burger, G., L. Forget, Y. Zhu, M. W. Gray and B. F. Lang (2003). "Unique mitochondrial genome architecture in unicellular relatives of animals." Proc Natl Acad Sci U S A **100**(3): 892-897.
Burger, G., M. W. Gray, L. Forget and B. F. Lang (2013). "Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists." Genome Biol Evol **5**(2): 418-438.
Burger, G., Y. Yan, P. Javadi and B. F. Lang (2009). "Group I-intron trans-splicing and mRNA editing in the mitochondria of placozoan animals." Trends Genet **25**(9): 381-386.
Cahan, P. and J. Kennell (2005). "Identification and distribution of sequences having similarity to mitochondrial plasmids in mitochondrial genomes of filamentous fungi." Molecular Genetics and Genomics **273**(6): 462-473.
Capella-Gutie´rrezS, Marcet-HoubenM,Gabaldo´nT(2012). "Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi". BMC Biol. 10:47.
Cech, T. R., N. K. Tanner, I. Tinoco, Jr., B. R. Weir, M. Zuker and P. S. Perlman (1983). "Secondary structure of the Tetrahymena ribosomal RNA intervening sequence: structural homology with fungal mitochondrial intervening sequences." Proc Natl Acad Sci U S A **80**(13): 3903-3907.
Cermakian, N., T. M. Ikeda, P. Miramontes, B. F. Lang, M. W. Gray and R. Cedergren (1997). "On the evolution of the single-subunit RNA polymerases." J Mol Evol **45**(6): 671-681.
Chan, D. C. (2006). "Mitochondria: dynamic organelles in disease, aging, and development." Cell **125**(7): 1241-1252.
Chen, H. and D. C. Chan (2009) "Mitochondrial dynamics--fusion, fission, movement, and mitophagy--in neurodegenerative diseases." (1460-2083 (Electronic)).
Chevalier, B., M. Turmel, C. Lemieux, R. J. Monnat, Jr. and B. L. Stoddard (2003). "Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI." J Mol Biol **329**(2): 253-269.
Cho, Y., Y.-L. Qiu, P. Kuhlman and J. D. Palmer (1998). "Explosive invasion of plant mitochondria by a group I intron." Proceedings of the National Academy of Sciences **95**(24): 14244-14249.
Clapp, J. P., J. P. W. Young, J. W. Merryweather and A. H. Fitter (1995). "Diversity of fungal symbionts in arbuscular mycorrhizas from a natural community." New Phytologist **130**(2): 259-265.
Colard, A., C. Angelard and I. R. Sanders (2011). "Genetic Exchange in an Arbuscular Mycorrhizal Fungus Results in Increased Rice Growth and Altered Mycorrhiza-Specific Gene Transcription." Applied and Environmental Microbiology **77**(18): 6510-6515.
Colleaux, L., L. d'Auriol, M. Betermier, G. Cottarel, A. Jacquier, F. Galibert and B. Dujon "Universal code equivalent of a yeast mitochondrial intron reading frame is expressed into E. coli as a specific double strand endonuclease." Cell **44**(4): 521-533.
Conway, D. J., J. M. Fanello C Fau - Lloyd, B. M. Lloyd Jm Fau - Al-Joubori, A. H. Al-Joubori Bm Fau - Baloch, S. D. Baloch Ah Fau - Somanath, C. Somanath Sd Fau - Roper, A. M. Roper C Fau - Oduola, B. Oduola Am Fau - Mulder, M. M. Mulder B Fau - Povoa, B. Povoa Mm Fau - Singh, A. W. Singh B Fau - Thomas and A. W. Thomas (2000). "Origin of Plasmodium falciparum malaria is traced by mitochondrial DNA." (0166-6851 (Print)).
Cordier, C., M. J. Pozo, J. M. Barea, S. Gianinazzi and V. Gianinazzi-Pearson (1998). "Cell Defense Responses Associated with Localized and Systemic Resistance to Phytophthora parasitica Induced in Tomato by an Arbuscular Mycorrhizal Fungus." Molecular Plant-Microbe Interactions **11**(10): 1017-1028.

Corradi, N., D. Croll, A. Colard, G. Kuhn, M. Ehinger and I. R. Sanders (2007). "Gene Copy Number Polymorphisms in an Arbuscular Mycorrhizal Fungal Population." Applied and Environmental Microbiology **73**(1): 366-369.
Croll, D., M. Giovannetti, A. M. Koch, C. Sbrana, M. Ehinger, P. J. Lammers and I. R. Sanders (2009). "Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus Glomus intraradices." New Phytologist **181**(4): 924-937.
Croll, D. and I. R. Sanders (2009). "Recombination in Glomus intraradices, a supposed ancient asexual arbuscular mycorrhizal fungus." BMC Evol Biol **9**: 13.
Dalgaard, J. Z., R. A. Garrett and M. Belfort (1993). "A site-specific endonuclease encoded by a typical archaeal intron." Proceedings of the National Academy of Sciences **90**(12): 5414-5417.
Datnoff, L. E., S. Nemec and K. Pernezny (1995). "Biological Control of Fusarium Crown and Root Rot of Tomato in Florida Using Trichoderma harzianum and Glomus intraradices." Biological Control **5**(3): 427-431.
Davila, J. I., M. P. Arrieta-Montiel, Y. Wamboldt, J. Cao, J. Hagmann, V. Shedge, Y. Z. Xu, D. Weigel and S. A. Mackenzie (2011). "Double-strand break repair processes drive evolution of the mitochondrial genome in Arabidopsis." BMC Biol **9**: 64.
de la Providencia, I. E., M. Nadimi, D. Beaudet, G. R. Morales and M. Hijri (2013). "Detection of a transient mitochondrial DNA heteroplasmy in the progeny of crossed genetically divergent isolates of arbuscular mycorrhizal fungi." New Phytologist **200**(1): 211-221.
Doner, L., x00E and G. card (1991). "Solubilization of gellan gels by chelation of cations." Biotechnology Techniques **5**(1): 25-28.
Drissi, R., F. Sor, J. Nosek and H. Fukuhara (1994). "Genes of the linear mitochondrial DNA of Williopsis mrakii: coding sequences for a maturase-like protein, a ribosomal protein VAR1 homologue, cytochrome oxidase subunit 2 and methionyl tRNA." Yeast **10**(3): 391-398.
Duchene, S., F. I. Archer, J. Vilstrup, S. Caballero and P. A. Morin (2011). "Mitogenome phylogenetics: the impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation." PLoS One **6**(11): e27138.
Dujon, B. (1980). "Sequence of the intron and flanking exons of the mitochondrial 21S rRNA gene of yeast strains having different alleles at the co and rib-1 loci." Cell **20**(1): 185-197.
Ebersberger, I., R. de Matos Simoes, A. Kupczok, M. Gube, E. Kothe, K. Voigt and A. von Haeseler (2012). "A consistent phylogenetic backbone for the fungi." Mol Biol Evol **29**(5): 1319-1334.
Eddy, S. (1992). Introns in the T-seven bacteriophages. Ph. D. Dissertation, University of Colorado.
Eddy S. (2008). HMMER website [Internet]. Ashburn (VA): HHMI Janelia Farm Research Campus. Available from: http://hmmer. janelia.org
Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Res **32**(5): 1792-1797.
Ehinger, M. O., D. Croll, A. M. Koch and I. R. Sanders (2012). "Significant genetic and phenotypic changes arising from clonal growth of a single spore of an arbuscular mycorrhizal fungus over multiple generations." New Phytologist **196**(3): 853-861.
Esch, H., B. Hundeshagen, H. Schneider-Poetsch and H. Bothe (1994). "Demonstration of abscisic acid in spores and hyphae of the arbuscular-mycorrhizal fungus Glomus and in the N2-fixing cyanobacterium Anabaena variabilis." Plant Science **99**(1): 9-16.
Felsenstein, J. (1978). "Cases in which Parsimony or Compatibility Methods Will be Positively Misleading." Systematic Zoology **27**(4): 401-410.
Fenn, J. D., H. Song, S. L. Cameron and M. F. Whiting (2008). "A preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data." Mol Phylogenet Evol **49**(1): 59-68.
Ferandon, C., S. E. K. Chatel, B. Castandet, M. Castroviejo and G. Barroso (2008). "The Agrocybe aegerita mitochondrial genome contains two inverted repeats of the nad4 gene arisen by

duplication on both sides of a linear plasmid integration site." Fungal Genetics and Biology **45**(3): 292-301.
Finn, R. D., J. Clements and S. R. Eddy (2011). "HMMER web server: interactive sequence similarity searching." Nucleic Acids Research **39**(suppl 2): W29-W37.
Formey, D., M. Moles, A. Haouy, B. Savelli, O. Bouchez, G. Becard and C. Roux (2012). "Comparative analysis of mitochondrial genomes of Rhizophagus irregularis - syn. Glomus irregulare - reveals a polymorphism induced by variability generating elements." New Phytologist **196**(4): 1217-1227.
Foury, F., T. Roganti, N. Lecrenier and B. Purnelle (1998). "The complete sequence of the mitochondrial genome of Saccharomyces cerevisiae." FEBS Lett **440**(3): 325-331.
Fricova, D., M. Valach, Z. Farkas, I. Pfeiffer, J. Kucsera, L. Tomaska and J. Nosek (2010). "The mitochondrial genome of the pathogenic yeast Candida subhashii: GC-rich linear DNA with a protein covalently attached to the 5′ termini." Microbiology **156**(7): 2153-2163.
Fu, C., P. Iyer, A. Herkal, J. Abdullah, A. Stout and S. J. Free (2011). "Identification and characterization of genes required for cell-to-cell fusion in Neurospora crassa." Eukaryot Cell **10**(8): 1100-1109.
Fukuhara, H., F. Sor, R. Drissi, N. Dinouel, I. Miyakawa, S. Rousset and A. M. Viola (1993). "Linear mitochondrial DNAs of yeasts: frequency of occurrence and general features." Mol Cell Biol **13**(4): 2309-2314.
G., H. (2011). Introns, mobile elements, and plasmids. Berlin (Germany), Springer Verlag.
Gandolfi, A., I. R. Sanders, V. Rossi and P. Menozzi (2003). "Evidence of recombination in putative ancient asexuals." Mol Biol Evol **20**(5): 754-761.
Garcia-Garrido, J. M. and J. A. Ocampo (2002). "Regulation of the plant defence response in arbuscular mycorrhizal symbiosis." J Exp Bot **53**(373): 1377-1386.
Gautheret, D. and A. Lambert "Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles." (0022-2836 (Print)).
Giovannetti, M., D. Azzolini and A. S. Citernesi (1999). "Anastomosis formation and nuclear and protoplasmic exchange in arbuscular mycorrhizal fungi." Applied and Environmental Microbiology **65**(12): 5571-5575.
Giovannetti, M., P. Fortuna, A. S. Citernesi, S. Morini and M. P. Nuti (2001). "The occurrence of anastomosis formation and nuclear exchange in intact arbuscular mycorrhizal networks." New Phytologist **151**(3): 717-724.
Giovannetti, M., C. Sbrana, P. Strani, M. Agnolucci, V. Rinaudo and L. Avio (2003). "Genetic diversity of isolates of Glomus mosseae from different geographic areas detected by vegetative compatibility testing and biochemical and molecular analysis." Applied and Environmental Microbiology **69**(1): 616-624.
Gobert, A., A. Gutmann B Fau - Taschner, M. Taschner A Fau - Gossringer, J. Gossringer M Fau - Holzmann, R. K. Holzmann J Fau - Hartmann, W. Hartmann Rk Fau - Rossmanith, P. Rossmanith W Fau - Giege and P. Giege "A single Arabidopsis organellar protein has RNase P activity." (1545-9985 (Electronic)).
Goldschmidt-Clermont, M., J. Choquet Y Fau - Girard-Bascou, F. Girard-Bascou J Fau - Michel, M. Michel F Fau - Schirmer-Rahire, J. D. Schirmer-Rahire M Fau - Rochaix and J. D. Rochaix "A small chloroplast RNA may be required for trans-splicing in Chlamydomonas reinhardtii." (0092-8674 (Print)).
Gollotte, A., D. Van Tuinen and D. Atkinson (2004). "Diversity of arbuscular mycorrhizal fungi colonising roots of the grass species Agrostis capillaris and Lolium perenne in a field experiment." Mycorrhiza **14**(2): 111-117.
Gray, M. W., B. F. Burger G Fau - Lang and B. F. Lang (1999). "Mitochondrial evolution." (0036-8075 (Print)).

Grewe, F., P. Viehoever, B. Weisshaar and V. Knoop (2009). "A trans-splicing group I intron and tRNA-hyperediting in the mitochondrial genome of the lycophyte Isoetes engelmannii." Nucleic Acids Res **37**(15): 5093-5104.
Griffiths, A. J. F. and X. Yang (1995). "Recombination between heterologous linear and circular mitochondrial plasmids in the fungus *Neurospora*." Molecular and General Genetics **249**(1): 25-36.
Halary, S., L. Daubois, Y. Terrat, S. Ellenberger, J. Wostemeyer and M. Hijri (2013). "Mating type gene homologues and putative sex pheromone-sensing pathway in arbuscular mycorrhizal fungi, a presumably asexual plant root symbiont." PLoS One **8**(11): e80729.
Halary, S., S.-B. Malik, L. Lildhar, C. H. Slamovits, M. Hijri and N. Corradi (2011). "Conserved Meiotic Machinery in Glomus spp., a Putatively Ancient Asexual Fungal Lineage." Genome Biology and Evolution **3**: 950-958.
Hausner, G. (2012). Introns, Mobile Elements, and Plasmids Organelle Genetics. C. E. Bullerwell, Springer Berlin Heidelberg**:** 329-357.
Havird, J. C. and S. R. Santos (2014). "Performance of single and concatenated sets of mitochondrial genes at inferring metazoan relationships relative to full mitogenome data." PLoS One **9**(1): e84080.
Hebert, P. D., S. Ratnasingham and J. R. deWaard (2003). "Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species." Proc Biol Sci **270 Suppl 1**: S96-99.
Hecht, J., F. Grewe and V. Knoop (2011). "Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of Selaginella moellendorffii mitochondria: the root of frequent plant mtDNA recombination in early tracheophytes." Genome Biology and Evolution **3**: 344-358.
Heinonen, T. Y., M. N. Schnare, P. G. Young and M. W. Gray (1987). "Rearranged coding segments, separated by a transfer RNA gene, specify the two parts of a discontinuous large subunit ribosomal RNA in Tetrahymena pyriformis mitochondria." J Biol Chem **262**(6): 2879-2887.
Helgason, T., A. H. Fitter and J. P. W. Young (1999). "Molecular diversity of arbuscular mycorrhizal fungi colonising Hyacinthoides non-scripta (bluebell) in a seminatural woodland." Molecular Ecology **8**(4): 659-666.
Helgason, T., I. J. Watson and J. P. Young (2003). "Phylogeny of the Glomerales and Diversisporales (fungi: Glomeromycota) from actin and elongation factor 1-alpha sequences." FEMS Microbiol Lett **229**(1): 127-132.
Hempel, S., C. Renker and F. Buscot (2007). "Differences in the species composition of arbuscular mycorrhizal fungi in spore, root and soil communities in a grassland ecosystem." Environtal Microbiology **9**(8): 1930-1938.
Hibbett, D. S., M. Binder, J. F. Bischoff, M. Blackwell, P. F. Cannon, O. E. Eriksson, S. Huhndorf, T. James, P. M. Kirk, R. Lucking, H. Thorsten Lumbsch, F. Lutzoni, P. B. Matheny, D. J. McLaughlin, M. J. Powell, S. Redhead, C. L. Schoch, J. W. Spatafora, J. A. Stalpers, R. Vilgalys, M. C. Aime, A. Aptroot, R. Bauer, D. Begerow, G. L. Benny, L. A. Castlebury, P. W. Crous, Y. C. Dai, W. Gams, D. M. Geiser, G. W. Griffith, C. Gueidan, D. L. Hawksworth, G. Hestmark, K. Hosaka, R. A. Humber, K. D. Hyde, J. E. Ironside, U. Koljalg, C. P. Kurtzman, K. H. Larsson, R. Lichtwardt, J. Longcore, J. Miadlikowska, A. Miller, J. M. Moncalvo, S. Mozley-Standridge, F. Oberwinkler, E. Parmasto, V. Reeb, J. D. Rogers, C. Roux, L. Ryvarden, J. P. Sampaio, A. Schussler, J. Sugiyama, R. G. Thorn, L. Tibell, W. A. Untereiner, C. Walker, Z. Wang, A. Weir, M. Weiss, M. M. White, K. Winka, Y. J. Yao and N. Zhang (2007). "A higher-level phylogenetic classification of the Fungi." Mycol Res **111**(Pt 5): 509-547.
Hijri, M. and I. R. Sanders (2005). "Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei." Nature **433**(7022): 160-163.
Hildebrandt, U., M. Kaldorf and H. Bothe (1999). "The Zinc Violet and its Colonization by Arbuscular Mycorrhizal Fungi." Journal of Plant Physiology **154**(5–6): 709-717.
Hildebrandt, U., F. Ouziad, F. J. Marner and H. Bothe (2006). "The bacterium Paenibacillus validus stimulates growth of the arbuscular mycorrhizal fungus Glomus intraradices up to the formation of fertile spores." FEMS Microbiol Lett **254**(2): 258-267.

Hnatova, M., Y. Gbelska, M. Obernauerova, V. Subikova and J. Subik (2003). "Cross-resistance to strobilurin fungicides in mitochondrial and nuclear mutants of Saccharomyces cerevisiae." Folia Microbiol (Praha) **48**(4): 496-500.
Hoekstra, R. F. (2000). "Evolutionary origin and consequences of uniparental mitochondrial inheritance." Hum Reprod **15 Suppl 2**: 102-111.
Holzmann, J., E. Frank P Fau - Loffler, K. L. Loffler E Fau - Bennett, C. Bennett Kl Fau - Gerner, W. Gerner C Fau - Rossmanith and W. Rossmanith "RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme." (1097-4172 (Electronic)).
Imhof, S. (1999). "Anatomy and mycotrophy of the achlorophyllous Afrothismia winkleri." New Phytologist **144**(3): 533-540.
Ismail, Y. and M. Hijri (2012). "Arbuscular mycorrhisation with Glomus irregulare induces expression of potato PR homologues genes in response to infection by Fusarium sambucinum." Functional Plant Biology **39**(3): 236-245.
Ismail, Y., S. McCormick and M. Hijri (2011). "A Fungal Symbiont of Plant-Roots Modulates Mycotoxin Gene Expression in the Pathogen Fusarium sambucinum." PLoS ONE **6**(3): e17990.
Ismail, Y., S. McCormick and M. Hijri (2013). "The arbuscular mycorrhizal fungus, Glomus irregulare, controls the mycotoxin production of Fusarium sambucinum in the pathogenesis of potato." FEMS Microbiol Lett **348**(1): 46-51.
Jacquier, A. and M. Rosbash (1986). "Efficient trans-splicing of a yeast mitochondrial RNA group II intron implicates a strong 5' exon-intron interaction." Science **234**(4780): 1099-1104.
James, T. Y., C. L. Kauff F Fau - Schoch, P. B. Schoch Cl Fau - Matheny, V. Matheny Pb Fau - Hofstetter, C. J. Hofstetter V Fau - Cox, G. Cox Cj Fau - Celio, C. Celio G Fau - Gueidan, E. Gueidan C Fau - Fraker, J. Fraker E Fau - Miadlikowska, H. T. Miadlikowska J Fau - Lumbsch, A. Lumbsch Ht Fau - Rauhut, V. Rauhut A Fau - Reeb, A. E. Reeb V Fau - Arnold, A. Arnold Ae Fau - Amtoft, J. E. Amtoft A Fau - Stajich, K. Stajich Je Fau - Hosaka, G.-H. Hosaka K Fau - Sung, D. Sung Gh Fau - Johnson, B. Johnson D Fau - O'Rourke, M. O'Rourke B Fau - Crockett, M. Crockett M Fau - Binder, J. M. Binder M Fau - Curtis, J. C. Curtis Jm Fau - Slot, Z. Slot Jc Fau - Wang, A. W. Wang Z Fau - Wilson, A. Wilson Aw Fau - Schussler, J. E. Schussler A Fau - Longcore, K. Longcore Je Fau - O'Donnell, S. O'Donnell K Fau - Mozley-Standridge, D. Mozley-Standridge S Fau - Porter, P. M. Porter D Fau - Letcher, M. J. Letcher Pm Fau - Powell, J. W. Powell Mj Fau - Taylor, M. M. Taylor Jw Fau - White, G. W. White Mm Fau - Griffith, D. R. Griffith Gw Fau - Davies, R. A. Davies Dr Fau - Humber, J. B. Humber Ra Fau - Morton, J. Morton Jb Fau - Sugiyama, A. Y. Sugiyama J Fau - Rossman, J. D. Rossman Ay Fau - Rogers, D. H. Rogers Jd Fau - Pfister, D. Pfister Dh Fau - Hewitt, K. Hewitt D Fau - Hansen, S. Hansen K Fau - Hambleton, R. A. Hambleton S Fau - Shoemaker, J. Shoemaker Ra Fau - Kohlmeyer, B. Kohlmeyer J Fau - Volkmann-Kohlmeyer, R. A. Volkmann-Kohlmeyer B Fau - Spotts, M. Spotts Ra Fau - Serdani, P. W. Serdani M Fau - Crous, K. W. Crous Pw Fau - Hughes, K. Hughes Kw Fau - Matsuura, E. Matsuura K Fau - Langer, G. Langer E Fau - Langer, W. A. Langer G Fau - Untereiner, R. Untereiner Wa Fau - Lucking, B. Lucking R Fau - Budel, D. M. Budel B Fau - Geiser, A. Geiser Dm Fau - Aptroot, P. Aptroot A Fau - Diederich, I. Diederich P Fau - Schmitt, M. Schmitt I Fau - Schultz, R. Schultz M Fau - Yahr, D. S. Yahr R Fau - Hibbett, F. Hibbett Ds Fau - Lutzoni, D. J. Lutzoni F Fau - McLaughlin, J. W. McLaughlin Dj Fau - Spatafora, R. Spatafora Jw Fau - Vilgalys and R. Vilgalys (2006). "Reconstructing the early evolution of Fungi using a six-gene phylogeny." (1476-4687 (Electronic)).
Jany, J. L. and T. E. Pawlowska (2010). "Multinucleate spores contribute to evolutionary longevity of asexual glomeromycota." Am Nat **175**(4): 424-435.
Jeffries, P., S. Gianinazzi, S. Perotto, K. Turnau and J.-M. Barea (2003). "The contribution of arbuscular mycorrhizal fungi in sustainable maintenance of plant health and soil fertility." Biology and Fertility of Soils **37**(1): 1-16.
Keeling, P. J. and C. H. Slamovits (2004). "Simplicity and complexity of microsporidian genomes." (1535-9778 (Print)).

Klassen, R. and F. Meinhardt (2007). Linear protein primed replicating plasmids in eukaryotic microbes. Microbial linear plasmids. Berlin, Germany, Springer**:** 188-226.
Knoop, V., M. Altwasser and A. Brennicke (1997). "A tripartite group II intron in mitochondria of an angiosperm plant." Mol Gen Genet **255**(3): 269-276.
Krüger, M., C. Krüger, C. Walker, H. Stockinger and A. Schüßler (2012). "Phylogenetic reference data for systematics and phylotaxonomy of arbuscular mycorrhizal fungi from phylum to species level." New Phytologist **193**(4): 970-984.
Kuhn, G., M. Hijri and I. R. Sanders (2001). "Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi." Nature **414**: 745-748.
Kuhn, G., M. Hijri and I. R. Sanders (2001). "Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi." Nature **414**(6865): 745-748.
Lambowitz, A. M. and S. Zimmerly (2004). "Mobile group II introns." Annu Rev Genet **38**: 1-35.
Lang, B. F. "The mitochondrial genome of the fission yeast Schizosaccharomyces pombe: highly homologous introns are inserted at the same position of the otherwise less conserved cox1 genes in Schizosaccharomyces pombe and Aspergillus nidulans." (0261-4189 (Print)).
Lang, B. F., G. Burger, C. J. O'Kelly, R. Cedergren, G. B. Golding, C. Lemieux, D. Sankoff, M. Turmel and M. W. Gray (1997). "An ancestral mitochondrial DNA resembling a eubacterial genome in miniature." Nature **387**(6632): 493-497.
Lang, B. F., R. Cedergren and M. W. Gray (1987). "The mitochondrial genome of the fission yeast, Schizosaccharomyces pombe. Sequence of the large-subunit ribosomal RNA gene, comparison of potential secondary structure in fungal mitochondrial large-subunit rRNAs and evolutionary considerations." Eur J Biochem **169**(3): 527-537.
Lang, B. F., M. W. Gray and G. Burger (1999). "Mitochondrial genome evolution and the origin of eukaryotes." Annu Rev Genet **33**: 351-397.
Lang, B. F. and M. Hijri (2009). "The complete Glomus intraradices mitochondrial genome sequence--a milestone in mycorrhizal research." New Phytol **183**(1): 3-6.
Lang, B. F., M. J. Laforest and G. Burger (2007). "Mitochondrial introns: a critical view." Trends Genet **23**(3): 119-125.
Lang, B. F., D. Lavrov, N. Beck and S. Steinberg (2012). Mitochondrial tRNA Structure, Identity, and Evolution of the Genetic Code. Organelle Genetics. C. E. Bullerwell, Springer Berlin Heidelberg**:** 431-474.
Lartillot, N., H. Brinkmann H Fau - Philippe and H. Philippe "Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model." (1471-2148 (Electronic)).
Lartillot, N. and H. Philippe (2004). "A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process." Mol Biol Evol **21**(6): 1095-1109.
Leake, J., D. Johnson, D. Donnelly, G. Muckle, L. Boddy and D. Read (2004). "Networks of power and influence: the role of mycorrhizal mycelium in controlling plant communities and agroecosystem functioning." Canadian Journal of Botany **82**(8): 1016-1045.
Lee, J., S. Lee and J. P. Young (2008). "Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi." FEMS Microbiol Ecol **65**(2): 339-349.
Lee, J. and J. P. Young (2009). "The mitochondrial genome sequence of the arbuscular mycorrhizal fungus Glomus intraradices isolate 494 and implications for the phylogenetic placement of Glomus." New Phytol **183**(1): 200-211.
Lesemann SS, S. S., Dunemann F, Deising HB (2006). "Mitochondrial heteroplasmy for the cytochrome b gene controls the level of strobilurin resistance in the apple powdery mildew fungus Podosphaera leucotricha (Ell. & Ev.) E.S. Salmon. ." Journal of Plant Diseases and Protection **113**: 259–266.
Lin, K., Limpens, E., Zhang, Z., Ivanov, S., Saunders, DGO., et al. (2014). "Single Nucleus Genome Sequencing Reveals High Similarity among Nuclei of an Endomycorrhizal Fungus." PLoS Genet 10(1).

Ling, F. and T. Shibata (2004). "Mhr1p-dependent concatemeric mitochondrial DNA formation for generating yeast mitochondrial homoplasmic cells." (1059-1524 (Print)).
Liu, Y., J. W. Leigh, H. Brinkmann, M. T. Cushion, N. Rodriguez-Ezpeleta, H. Philippe and B. F. Lang (2009). "Phylogenomic analyses support the monophyly of Taphrinomycotina, including Schizosaccharomyces fission yeasts." Mol Biol Evol **26**(1): 27-34.
Liu, Y., E. T. Steenkamp, H. Brinkmann, L. Forget, H. Philippe and B. F. Lang (2009). "Phylogenomic analyses predict sistergroup relationship of nucleariids and fungi and paraphyly of zygomycetes with significant support." BMC Evol Biol **9**: 272.
Lockshon, D., S. G. Zweifel, L. L. Freeman-Cook, H. E. Lorimer, B. J. Brewer and W. L. Fangman (1995). "A role for recombination junctions in the segregation of mitochondrial DNA in yeast." Cell **81**(6): 947-955.
Lohse, M., O. Drechsel, S. Kahlau and R. Bock (2013). "OrganellarGenomeDRAW--a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets." Nucleic Acids Res **41**(Web Server issue): W575-581.
Lynch, M. (1996). "Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes." Molecular Biology and Evolution **13**(1): 209-220.
MacAlpine, D. M., P. S. Perlman and R. A. Butow (1998). "The high mobility group protein Abf2p influences the level of yeast mitochondrial DNA recombination intermediates in vivo." Proc Natl Acad Sci U S A **95**(12): 6739-6743.
Mamirova, L., K. Popadin and M. S. Gelfand (2007). "Purifying selection in mitochondria, free-living and obligate intracellular proteobacteria." BMC Evol Biol **7**: 17.
Marechal, A. and N. Brisson (2010). "Recombination and the maintenance of plant organelle genome stability." New Phytol **186**(2): 299-317.
Margulis, L. (1971). Origin of Eukaryotic Cells, Yale University Press.
Marleau, J., Y. Dalpe, M. St-Arnaud and M. Hijri (2011). "Spore development and nuclear inheritance in arbuscular mycorrhizal fungi." BMC Evol Biol **11**: 51.
Martin, F., V. Gianinazzi-Pearson, M. Hijri, P. Lammers, N. Requena, I. R. Sanders, Y. Shachar-Hill, H. Shapiro, G. A. Tuskan and J. P. Young (2008). "The long hard road to a completed Glomus intraradices genome." New Phytol **180**(4): 747-750.
Martin, F. N. (1995). "Linear mitochondrial genome organization in vivo in the genus Pythium." Curr Genet **28**(3): 225-234.
Mathews, D. H., M. D. Disney, J. L. Childs, S. J. Schroeder, M. Zuker and D. H. Turner (2004). "Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure." Proc Natl Acad Sci U S A **101**(19): 7287-7292.
Michel, F., A. Jacquier and B. Dujon (1982). "Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure." Biochimie **64**(10): 867-881.
Miller, M. A., W. Pfeiffer and T. Schwartz (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Gateway Computing Environments Workshop (GCE), 2010.
Moritz, C. and W. M. Brown "Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards." (0027-8424 (Print)).
Morton JB, B. G. (1990). "Revised classification of arbuscular mycorrhizal fungi (Zygomycetes): a new order, Glomales, two new suborders, Glomineae and Gigasporineae, and two families, Acaulosporaceae and Gigasporaceae, with an emendation of Glomaceae." Mycotaxon **37**: 471–491.
Msiska, Z. and J. B. Morton (2009). "Phylogenetic analysis of the Glomeromycota by partial beta-tubulin gene sequences." Mycorrhiza **19**(4): 247-254.
Muramatsu, T., S. Yokoyama, N. Horie, A. Matsuda, T. Ueda, Z. Yamaizumi, Y. Kuchino, S. Nishimura and T. Miyazawa (1988). "A novel lysine-substituted nucleoside in the first position of the anticodon of minor isoleucine tRNA from Escherichia coli." J Biol Chem **263**(19): 9261-9267.

Nadimi, M., D. Beaudet, L. Forget, M. Hijri and B. F. Lang (2012). "Group I intron-mediated trans-splicing in mitochondria of Gigaspora rosea, and a robust phylogenetic affiliation of arbuscular mycorrhizal fungi with Mortierellales." Molecular Biology and Evolution.
Ni, M., M. Feretzaki, S. Sun, X. Wang and J. Heitman (2011). "Sex in fungi." Annu Rev Genet **45**: 405-430.
Nosek, J. and L. Tomaska (2003). "Mitochondrial genome diversity: evolution of the molecular architecture and replication strategy." Curr Genet **44**(2): 73-84.
Paquin, B., M. J. Laforest and B. F. Lang (1994). "Interspecific transfer of mitochondrial genes in fungi and creation of a homologous hybrid gene." Proceedings of the National Academy of Sciences **91**(25): 11807-11810.
Parniske, M. (2008). "Arbuscular mycorrhiza: the mother of plant root endosymbioses." Nat Rev Microbiol **6**(10): 763-775.
Pawlowska, T. E. and J. W. Taylor (2004). "Organization of genetic variation in individuals of arbuscular mycorrhizal fungi." Nature **427**(6976): 733-737.
Pawlowska, T. E. and J. W. Taylor (2005). "Arbuscular mycorrhizal fungi: Hyphal fusion and multigenomic structure (reply)." Nature **433**(7022): E4-E4.
Pelin, A., J. F. Pombert, A. Salvioli, L. Bonen, P. Bonfante and N. Corradi (2012). "The mitochondrial genome of the arbuscular mycorrhizal fungus Gigaspora margarita reveals two unsuspected trans-splicing events of group I introns." New Phytologist **194**(3): 836-845.
Perrin, A., M. Buckle and B. Dujon (1993). "Asymmetrical recognition and activity of the I-SceI endonuclease on its site and on intron-exon junctions." EMBO J **12**(7): 2939-2947.
Phillips, J. M. and D. S. Hayman (1970). "Improved procedures for clearing roots and staining parasitic and vesicular-arbuscular mycorrhizal fungi for rapid assessment of infection." Transactions of the British Mycological Society **55**(1): 158-IN118.
Pivato, B., S. Mazurier, P. Lemanceau, S. Siblot, G. Berta, C. Mougel and D. van Tuinen (2007). "Medicago species affect the community composition of arbuscular mycorrhizal fungi associated with roots." New Phytol **176**(1): 197-210.
Pombert, J. F. and P. J. Keeling (2010). "The mitochondrial genome of the entomoparasitic green alga helicosporidium." PLoS One **5**(1): e8954.
Pozo, M. J., C. Cordier, E. Dumas-Gaudot, S. Gianinazzi, J. M. Barea and C. Azcón-Aguilar (2002). "Localized versus systemic effect of arbuscular mycorrhizal fungi on defence responses to Phytophthora infection in tomato plants." Journal of Experimental Botany **53**(368): 525-534.
Purin, S. and J. B. Morton (2012). "Anastomosis behavior differs between asymbiotic and symbiotic hyphae of Rhizophagus clarus." (0027-5514 (Print)).
Raab, P. A., A. Brennwald and D. Redecker (2005). "Mitochondrial large ribosomal subunit sequences are homogeneous within isolates of Glomus (arbuscular mycorrhizal fungi, Glomeromycota)." Mycol Res **109**(Pt 12): 1315-1322.
Rand, D. M. (2001). "THE UNITS OF SELECTION ON MITOCHONDRIAL DNA." Annual Review of Ecology and Systematics **32**(1): 415-448.
Ratnasingham, S. and P. D. Hebert (2007). "bold: The Barcode of Life Data System (http://www.barcodinglife.org)." Mol Ecol Notes **7**(3): 355-364.
Rattray, A. J., C. B. McGill, B. K. Shafer and J. N. Strathern (2001). "Fidelity of mitotic double-strand-break repair in Saccharomyces cerevisiae: a role for SAE2/COM1." Genetics **158**(1): 109-122.
Rattray, A. J., B. K. Shafer, B. Neelam and J. N. Strathern (2005). "A mechanism of palindromic gene amplification in Saccharomyces cerevisiae." Genes Dev **19**(11): 1390-1399.
Ravnskov, S. and I. Jakobsen (1995). "Functional compatibility in arbuscular mycorrhizas measured as hyphal P transport to the plant." New Phytologist **129**(4): 611-618.
Rayner, A. D. M. (1991). "The Challenge of the Individualistic Mycelium." Mycologia **83**(1): 48-71.
Redecker, D. (2002). "Molecular identification and phylogeny of arbuscular mycorrhizal fungi." Plant and Soil **244**(1-2): 67-73.

Redecker, D., R. Kodner and L. E. Graham (2000). "Glomalean fungi from the Ordovician." Science **289**(5486): 1920-1921.
Redecker, D., J. B. Morton and T. D. Bruns (2000). "Ancestral lineages of arbuscular mycorrhizal fungi (Glomales)." Mol Phylogenet Evol **14**(2): 276-284.
Remy, W., T. N. Taylor, H. Hass and H. Kerp (1994). "Four hundred-million-year-old vesicular arbuscular mycorrhizae." Proc Natl Acad Sci U S A **91**(25): 11841-11843.
Requena, N., M. Breuninger, P. Franken and A. Ocon (2003). "Symbiotic status, phosphate, and sucrose regulate the expression of two plasma membrane H+-ATPase genes from the mycorrhizal fungus Glomus mosseae." Plant Physiol **132**(3): 1540-1549.
Rillig, M. C. and D. L. Mummey "Mycorrhizas and soil structure." (0028-646X (Print)).
Robison, M. M. and D. J. Wolyn (2005). "A mitochondrial plasmid and plasmid-like RNA and DNA polymerases encoded within the mitochondrial genome of carrot *Daucus carota*." Current Genetics **47**(1): 57-66.
Roe, A. D. and F. A. Sperling (2007). "Patterns of evolution of mitochondrial cytochrome c oxidase I and II DNA and implications for DNA barcoding." Mol Phylogenet Evol **44**(1): 325-345.
Romero, D. and R. Palacios (1997). "Gene amplification and genomic plasticity in prokaryotes." Annu Rev Genet **31**: 91-111.
Rosendahl, S., P. McGee and J. B. Morton (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." Mol Ecol **18**(20): 4316-4329.
Sanchez-Puerta, M. V., Y. Cho, J. P. Mower, A. J. Alverson and J. D. Palmer (2008). "Frequent, Phylogenetically Local Horizontal Transfer of the cox1 Group I Intron in Flowering Plant Mitochondria." Molecular Biology and Evolution **25**(8): 1762-1777.
Sanders, Ian R. (2011). "Fungal Sex: Meiosis Machinery in Ancient Symbiotic Fungi." Current biology **21**(21): R896-R897.
Sanders, I. R., M. Alt, K. Groppe, T. Boller and A. Wiemken (1995). "Identification of ribosomal DNA polymorphisms among and within spores of the Glomales: application to studies on the genetic diversity of arbuscular mycorrhizal fungal communities." New Phytologist **130**(3): 419-427.
Saupe, S. J. (2000). "Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes." Microbiol Mol Biol Rev **64**(3): 489-502.
Schmelzer, C. and R. J. Schweyen (1986). "Self-splicing of group II introns in vitro: mapping of the branch point and mutational inhibition of lariat formation." Cell **46**(4): 557-565.
Schnare, M. N. and M. W. Gray (1990). "Sixteen discrete RNA components in the cytoplasmic ribosome of Euglena gracilis." J Mol Biol **215**(1): 73-83.
Schnare, M. N., T. Y. Heinonen, P. G. Young and M. W. Gray (1986). "A discontinuous small subunit ribosomal RNA in Tetrahymena pyriformis mitochondria." J Biol Chem **261**(11): 5187-5193.
Schoch, C. L. and K. A. Seifert (2012). "Reply to Kiss: Internal transcribed spacer (ITS) remains the best candidate as a universal DNA barcode marker for Fungi despite imperfections." Proceedings of the National Academy of Sciences **109**(27): E1812-E1812.
Schüßler A, W. C. (2010). The Glomeromycota: a species list with new families and new genera.
. Edinburgh & Kew, UK: The Royal Botanic Garden; Munich, Germany: Botanische Staatssammlung Munich; Oregon, USA: Oregon State University. URL: http://www.amf-phylogeny.com. .
Schüßler, A., D. Schwarzott and C. Walker (2001). "A new fungal phylum, the Glomeromycota: phylogeny and evolution." Mycological Research **105**(12): 1413-1421.
Schwarzott, D., A. Walker C Fau - Schussler and A. Schussler "Glomus, the largest genus of the arbuscular mycorrhizal fungi (Glomales), is nonmonophyletic." (1055-7903 (Print)).
Seif, E., A. Cadieux and B. F. Lang (2006). "Hybrid E. coli--Mitochondrial ribonuclease P RNAs are catalytically active." RNA **12**(9): 1661-1670.

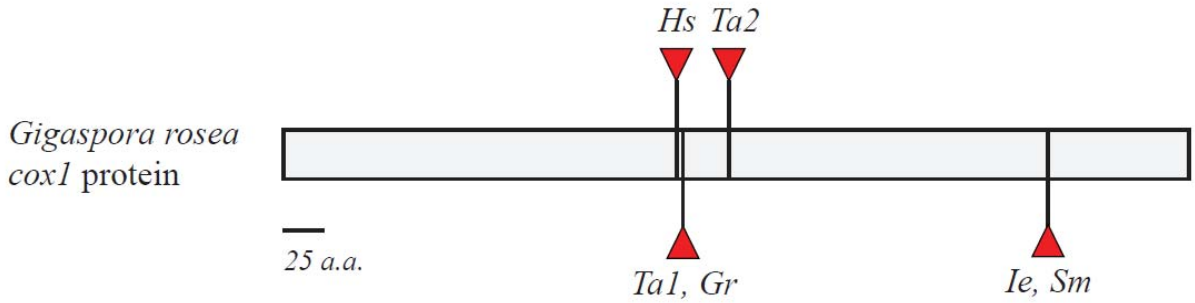
Seif, E., J. Leigh, Y. Liu, I. Roewer, L. Forget and B. F. Lang (2005). "Comparative mitochondrial genomics in zygomycetes: bacteria-like RNase P RNAs, mobile elements and a close source of the group I intron invasion in angiosperms." Nucleic Acids Res **33**(2): 734-744.
Shao, R., X. Q. Zhu, S. C. Barker and K. Herd (2012). "Evolution of extensively fragmented mitochondrial genomes in the lice of humans." Genome Biol Evol **4**(11): 1088-1101.
Slater, G. S. and E. Birney (2005). "Automated generation of heuristics for biological sequence comparison." BMC Bioinformatics **6**: 31.
Sloan, D. B., A. J. Alverson, J. P. Chuckalovcak, M. Wu, D. E. McCauley, J. D. Palmer and D. R. Taylor (2012). "Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates." PLoS Biol **10**(1): e1001241.
Smith, G. R. (1988). "Homologous recombination in procaryotes." Microbiol Rev **52**(1): 1-28.
Smith, M. L., L. C. Duchesne, J. N. Bruhn and J. B. Anderson (1990). "Mitochondrial genetics in a natural population of the plant pathogen armillaria." Genetics **126**(3): 575-582.
Smith, S. and D. Read, Eds. (2008). Mycorrhizal Symbiosis. Cambridge, UK: Academic press.
Smith SE, R. D. (1997). Mycorrhizal Symbiosis(Second Edition). London, Academic Press.
Sokolski, S., Y. Dalpe and Y. Piche (2011). "Phosphate transporter genes as reliable gene markers for the identification and discrimination of arbuscular mycorrhizal fungi in the genus glomus." Appl Environ Microbiol **77**(5): 1888-1891.
Sokolski, S., Y. Dalpé, S. Séguin, D. Khasa, C. A. Lévesque and Y. Piché (2010). "Conspecificity of DAOM 197198, the model arbuscular mycorrhizal fungus, with Glomus irregulare: molecular evidence with three protein-encoding genes." Botany **88**(9): 829-838.
St-Arnaud, M. and V. Vujanovik (2007). "Effects of Arbuscular Mycorrhizal Fungi on Plant Diseases and Pests " Mycorrhizae in Crop Production: Applying knowledge, Haworth Press, Binghampton, NY.
Stamatakis, A. (2006). "RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models." Bioinformatics **22**(21): 2688-2690.
Stamatakis, A., P. Hoover and J. Rougemont (2008). "A rapid bootstrap algorithm for the RAxML Web servers." Syst Biol **57**(5): 758-771.
Stockinger, H., M. Kruger and A. Schussler (2010). "DNA barcoding of arbuscular mycorrhizal fungi." New Phytol **187**(2): 461-474.
Stockinger, H., C. Walker and A. Schussler (2009). "'Glomus intraradices DAOM197198', a model fungus in arbuscular mycorrhiza research, is not Glomus intraradices." New Phytol **183**(4): 1176-1187.
Strack, D., T. Fester, B. Hause, W. Schliemann and M. H. Walter (2003). "Arbuscular mycorrhiza: biological, chemical, and molecular aspects." J Chem Ecol **29**(9): 1955-1979.
Stukenbrock, E. H. and S. Rosendahl (2005). "Development and amplification of multiple co-dominant genetic markers from single spores of arbuscular mycorrhizal fungi by nested multiplex PCR." Fungal Genet Biol **42**(1): 73-80.
Subramanian, K. S. and C. Charest (1999). "Acquisition of N by external hyphae of an arbuscular mycorrhizal fungus and its impact on physiological responses in maize under drought-stressed and well-watered conditions." Mycorrhiza **9**(2): 69-75.
Subramanian KS, Charest C 1999. Acquisition of N by external hyphae of an arbuscular mycorrhizal fungus and its impact on physiological responses in maize under drought-stressed and well-watered conditions. Mycorrhiza 9: 69-75. doi: Doi 10.1007/S005720050289
Suga, K., Y. Mark Welch Db Fau - Tanaka, Y. Tanaka Y Fau - Sakakura, A. Sakakura Y Fau - Hagiwara and A. Hagiwara "Two circular chromosomes of unequal copy number make up the mitochondrial genome of the rotifer Brachionus plicatilis." (1537-1719 (Electronic)).
Taylor, D. L. and T. D. Bruns (1999). "Community structure of ectomycorrhizal fungi in a Pinus muricata forest: minimal overlap between the mature forest and resistant propagule communities." Molecular Ecology **8**(11): 1837-1850.

Taylor, D. R., C. Zeyl and E. Cooke (2002). "Conflicting levels of selection in the accumulation of mitochondrial defects in Saccharomyces cerevisiae." Proc Natl Acad Sci U S A **99**(6): 3690-3694.
Thiery, O., B. Borstler, K. Ineichen and D. Redecker (2010). "Evolutionary dynamics of introns and homing endonuclease ORFs in a region of the large subunit of the mitochondrial rRNA in Glomus species (arbuscular mycorrhizal fungi, Glomeromycota)." Mol Phylogenet Evol **55**(2): 599-610.
Thiéry, O., B. Börstler, K. Ineichen and D. Redecker (2010). "Evolutionary dynamics of introns and homing endonuclease ORFs in a region of the large subunit of the mitochondrial rRNA in Glomus species (arbuscular mycorrhizal fungi, Glomeromycota)." Molecular Phylogenetics and Evolution **55**(2): 599-610.
Tisserant, E., A. Kohler, P. Dozolme-Seddas, R. Balestrini, K. Benabdellah, A. Colard, D. Croll, C. Da Silva, S. K. Gomez, R. Koul, N. Ferrol, V. Fiorilli, D. Formey, P. Franken, N. Helber, M. Hijri, L. Lanfranco, E. Lindquist, Y. Liu, M. Malbreil, E. Morin, J. Poulain, H. Shapiro, D. van Tuinen, A. Waschke, C. Azcón-Aguilar, G. Bécard, P. Bonfante, M. J. Harrison, H. Küster, P. Lammers, U. Paszkowski, N. Requena, S. A. Rensing, C. Roux, I. R. Sanders, Y. Shachar-Hill, G. Tuskan, J. P. W. Young, V. Gianinazzi-Pearson and F. Martin (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., M. Malbreil, A. Kuo, A. Kohler, A. Symeonidi, R. Balestrini, P. Charron, N. Duensing, N. Frei dit Frey, V. Gianinazzi-Pearson, L. B. Gilbert, Y. Handa, J. R. Herr, M. Hijri, R. Koul, M. Kawaguchi, F. Krajinski, P. J. Lammers, F. G. Masclaux, C. Murat, E. Morin, S. Ndikumana, M. Pagni, D. Petitpierre, N. Requena, P. Rosikiewicz, R. Riley, K. Saito, H. San Clemente, H. Shapiro, D. van Tuinen, G. Becard, P. Bonfante, U. Paszkowski, Y. Y. Shachar-Hill, G. A. Tuskan, J. P. Young, I. R. Sanders, B. Henrissat, S. A. Rensing, I. V. Grigoriev, N. Corradi, C. Roux and F. Martin (2013). "Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis." Proc Natl Acad Sci U S A **110**(50): 20117-20122.
Tuomi, J., M.-M. Kytöviita and R. Härdling (2001). "Cost efficiency of nutrient acquisition and the advantage of mycorrhizal symbiosis for the host plant." Oikos **92**(1): 62-70.
Untergasser, A., T. Cutcutache I Fau - Koressaar, J. Koressaar T Fau - Ye, B. C. Ye J Fau - Faircloth, M. Faircloth Bc Fau - Remm, S. G. Remm M Fau - Rozen and S. G. Rozen "Primer3--new capabilities and interfaces." (1362-4962 (Electronic)).
Valach, M., Z. Farkas, D. Fricova, J. Kovac, B. Brejova, T. Vinar, I. Pfeiffer, J. Kucsera, L. Tomaska, B. F. Lang and J. Nosek (2011). "Evolution of linear chromosomes and multipartite genomes in yeast mitochondria." Nucleic Acids Res **39**(10): 4202-4219.
van der Heijden, M. G., N. M. Bardgett Rd Fau - van Straalen and N. M. van Straalen (2008). "The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems." (1461-0248 (Electronic)).
van der Heijden, M. G., R. Streitwolf-Engel, R. Riedl, S. Siegrist, A. Neudecker, K. Ineichen, T. Boller, A. Wiemken and I. R. Sanders (2006). "The mycorrhizal contribution to plant productivity, plant nutrition and soil structure in experimental grassland." New Phytol **172**(4): 739-752.
van der Veen, R., A. C. Arnberg, G. van der Horst, L. Bonen, H. F. Tabak and L. A. Grivell (1986). "Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro." Cell **44**(2): 225-234.
VanKuren, N. W., H. C. den Bakker, J. B. Morton and T. E. Pawlowska (2013). "Ribosomal RNA gene diversity, effective population size, and evolutionary longevity in asexual Glomeromycota." Evolution **67**(1): 207-224.
Vaughn, J. C., M. T. Mason, G. L. Sper-Whitis, P. Kuhlman and J. D. Palmer (1995). "Fungal origin by horizontal transfer of a plant mitochondrial group I intron in the chimeric CoxI gene of Peperomia." J Mol Evol **41**(5): 563-572.

Vierula, P. J. and H. Bertrand (1992). "A deletion derivative of the kalilo senescence plasmid forms hairpin and duplex DNA structures in the mitochondria of Neurospora." Mol Gen Genet **234**(3): 361-368.
Walker, C., Sch, uuml, szlig and A. ler (2004). "NOMENCLATURAL CLARIFICATIONS AND NEW TAXA IN THE GLOMEROMYCOTA." Mycological Research **108**(09): 981-982.
Walker, C., M. Vestberg, F. Demircik, H. Stockinger, M. Saito, H. Sawaki, I. Nishmura and A. Schussler (2007). "Molecular phylogeny and new taxa in the Archaeosporales (Glomeromycota): Ambispora fennica gen. sp. nov., Ambisporaceae fam. nov., and emendation of Archaeospora and Archaeosporaceae." Mycol Res **111**(Pt 2): 137-153.
Wang, B. and Y. L. Qiu (2006). "Phylogenetic distribution and evolution of mycorrhizas in land plants." Mycorrhiza **16**(5): 299-363.
Waring, R. B., R. W. Davies, C. Scazzocchio and T. A. Brown (1982). "Internal structure of a mitochondrial intron of Aspergillus nidulans." Proc Natl Acad Sci U S A **79**(20): 6332-6336.
Weber, F., J. H. Dietrich A Fau - Weil, L. Weil Jh Fau - Marechal-Drouard and L. Marechal-Drouard "A potato mitochondrial isoleucine tRNA is coded for by a mitochondrial gene possessing a methionine anticodon." (0305-1048 (Print)).
White, D. J., J. N. Wolff, M. Pierson and N. J. Gemmell (2008). "Revealing the hidden complexities of mtDNA inheritance." Mol Ecol **17**(23): 4925-4942.
White, M. F. and D. M. Lilley (1997). "Characterization of a Holliday junction-resolving enzyme from Schizosaccharomyces pombe." Mol Cell Biol **17**(11): 6465-6471.
White, T., T. Bruns, S. Lee and J. Taylor (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR Protocols: A Guide to Methods and Applications. M. Innis, D. Gelfand, J. Shinsky and T. White, Academic Press**:** 315-322.
Wilson, R. J. and D. H. Williamson (1997). "Extrachromosomal DNA in the Apicomplexa." Microbiol Mol Biol Rev **61**(1): 1-16.
Wubet, T., I. Weiss M Fau - Kottke, D. Kottke I Fau - Teketay, F. Teketay D Fau - Oberwinkler and F. Oberwinkler "Phylogenetic analysis of nuclear small subunit rDNA sequences suggests that the endangered African Pencil Cedar, Juniperus procera, is associated with distinct members of Glomeraceae." (0953-7562 (Print)).
Wubet, T., M. Weiß, I. Kottke and F. Oberwinkler (2003). "Morphology and molecular diversity of arbuscular mycorrhizal fungi in wild and cultivated yew (Taxus baccata)." Canadian Journal of Botany **81**(3): 255-266.
Yang, H., Y. Zang, Y. Yuan, J. Tang and X. Chen (2012). "Selectivity by host plants affects the distribution of arbuscular mycorrhizal fungi: evidence from ITS rDNA sequence metadata." BMC Evol Biol **12**: 50.
Yang, X. and A. J. Griffiths (1993). "Male transmission of linear plasmids and mitochondrial DNA in the fungus Neurospora." Genetics **134**(4): 1055-1062.
Zhang, X., B. H. Ren, S. L. Wu, Y. Q. Sun, G. Lin and B. D. Chen (2014). "Arbuscular mycorrhizal symbiosis influences arsenic accumulation and speciation in Medicago truncatula L. in arsenic-contaminated soil." Chemosphere **119C**: 224-230.

Annex 1 : Supplementary Information (Chapter 2)

Group I intron-mediated trans-splicing in mitochondria of *Gigaspora rosea*, and a robust phylogenetic affiliation of arbuscular mycorrhizal fungi with *Mortierellales*



Supplementary Figure S2.1. Comparison of the *Helicosporidium* sp. (Hs), *Trichoplax adhaerens* (Ta), *Gigaspora rosea* (Gr), *Isoetes engelmannii* (Ie) and *Selaginella moellendorffii* (Sm) *cox1* group I intron trans-splicing positions represented on the *Gigaspora rosea cox1* protein sequence. (Hecht, Grewe et al. , Burger, Yan et al. 2009, Grewe, Viehoever et al. 2009, Pombert and Keeling 2010)

Supplementary Figure S2.2. Predicted UUG translation initiation in *Gigaspora* and *Mortierella*. Alignments of three mtDNA-encoded proteins are shown to demonstrate predictions of UUG translation initiation (with M standing for methionine, labeled red). The first dozen amino acid positions other than the predicted UUG methionine codons are color-coded according to amino acid family (mauve, P,A,G,S,T; green, M,L,I,V; orange, K,R,H; and blue, F, Y), to facilitate visual recognition of sequence similarity. Note that *cob* in *Mortierella* (the alleged phylogenetic neighbor of Glomeromycota) is predicted to start with a UUG initiation codon as noted previously (Seif, Leigh et al. 2005), yet in this species restricted to only one out of 14 protein coding genes.

Annex 2 : Supplementary Information (Chapter 3)

Rapid mitochondrial genome evolution through invasion of mobile elements in two closely related species of arbuscular mycorrhizal fungi

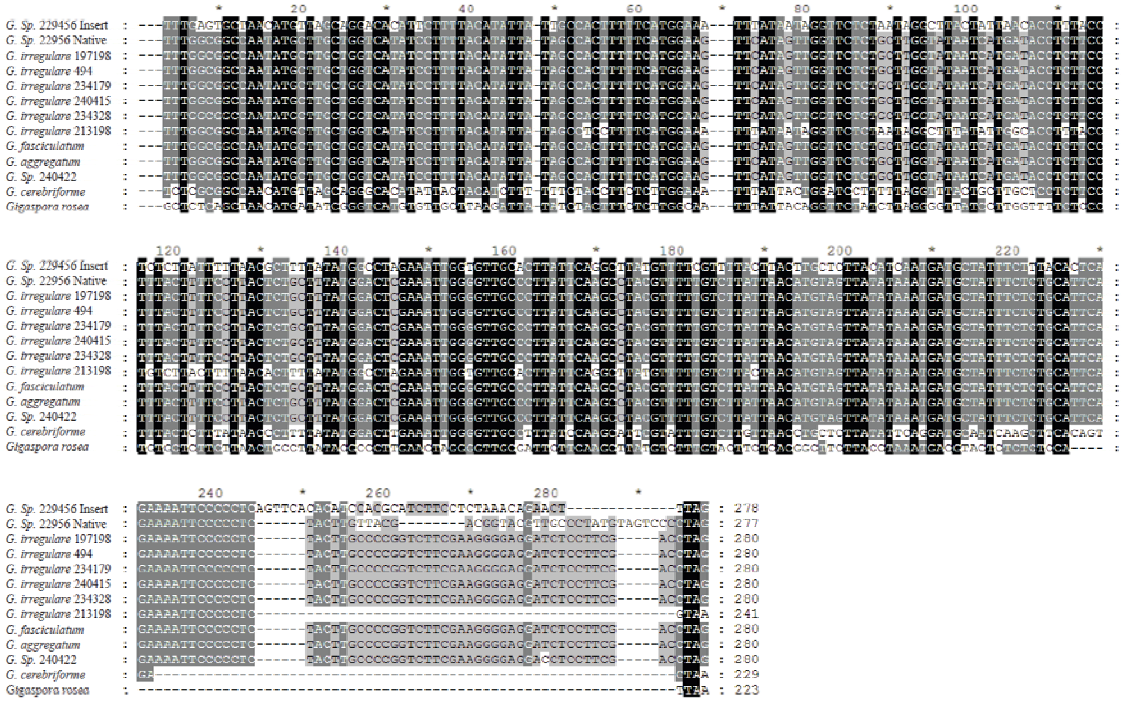


Figure S3.1. Multiple DNA sequence alignment of numerous AMF representatives of the *atp6* native C-terminals along with the *Glomus sp. 229456* putative foreign inserted C*-terminal.

Table S3.1. Sequence identity matrix of the atp6 native C-terminals along with the *Glomus* sp. 229456 putative foreign inserted C*-terminal.

Seq->	Gsp229456 Insert	Gsp229456 Native	Gi197198	Gi494	Gi234179	Gi240415	Gi234328	Gsp213198	fascicula	aggregatum	Gsp240422	cerebri	G._rosea
Gsp229456 Insert	ID	63.5%	68.1%	68.1%	68.1%	68.1%	68.1%	74.1%	68.1%	68.1%	68.1%	56.1%	50.3%
Gsp229456 Native	63.5%	ID	91.2%	91.2%	91.2%	91.2%	91.2%	72.5%	91.2%	91.2%	90.8%	56.6%	49.0%
Gi197198	68.1%	91.2%	ID	100.0%	100.0%	100.0%	100.0%	71.7%	100.0%	100.0%	99.6%	56.0%	48.5%
Gi494	68.1%	91.2%	100.0%	ID	100.0%	100.0%	100.0%	71.7%	100.0%	100.0%	99.6%	56.0%	48.5%
Gi234179	68.1%	91.2%	100.0%	100.0%	ID	100.0%	100.0%	71.7%	100.0%	100.0%	99.6%	56.0%	48.5%
Gi240415	68.1%	91.2%	100.0%	100.0%	100.0%	ID	100.0%	71.7%	100.0%	100.0%	99.6%	56.0%	48.5%
Gi234328	68.1%	91.2%	100.0%	100.0%	100.0%	100.0%	ID	71.7%	100.0%	100.0%	99.6%	56.0%	48.5%
Gsp213198	74.1%	72.5%	71.7%	71.7%	71.7%	71.7%	71.7%	ID	71.7%	71.7%	71.7%	63.9%	59.3%
fascicula	68.1%	91.2%	100.0%	100.0%	100.0%	100.0%	100.0%	71.7%	ID	100.0%	99.6%	56.0%	48.5%
aggregatum	68.1%	91.2%	100.0%	100.0%	100.0%	100.0%	100.0%	71.7%	100.0%	ID	99.6%	56.0%	48.5%
Gsp240422	68.1%	90.8%	99.6%	99.6%	99.6%	99.6%	99.6%	71.7%	99.6%	99.6%	ID	56.0%	48.5%
cerebri	56.1%	56.6%	56.0%	56.0%	56.0%	56.0%	56.0%	63.9%	56.0%	56.0%	56.0%	ID	63.7%
G._rosea	50.3%	49.0%	48.5%	48.5%	48.5%	48.5%	48.5%	59.3%	48.5%	48.5%	48.5%	63.7%	ID

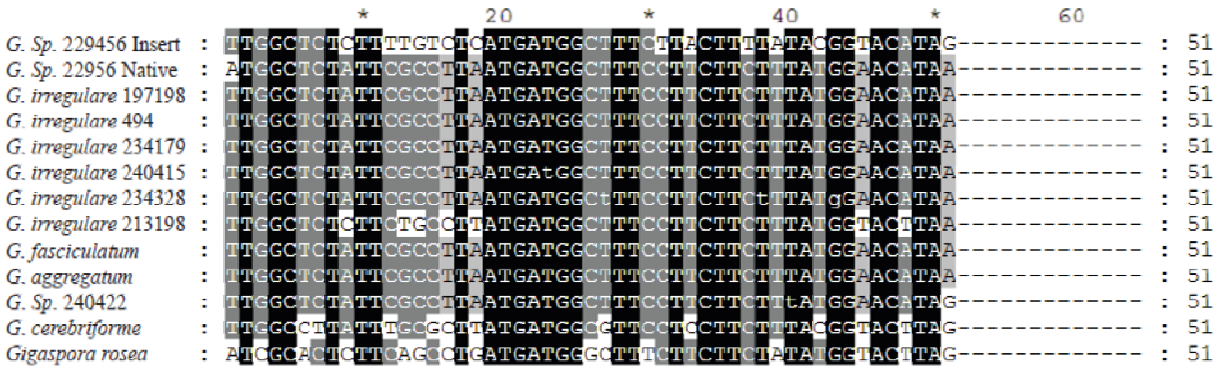


Figure S3.2. Multiple DNA sequence alignment of numerous AMF representatives of the *atp9* native C-terminals along with the *Glomus sp. 229456* putative foreign inserted C*- terminal.

Table S3.2. Sequence identity matrix of the atp9 native C-terminals along with the *Glomus sp.* 229456 putative foreign inserted C*-terminal.

Seq->	Gsp229456 Insert	Gsp229456 Native	Gi197198	Gi494	Gi234179	Gi240415	Gi234328	Gi213198	fascicula	aggregatum	Gsp240422	cerebri	G._rosea
Gsp229456 Insert	ID	70.5%	72.5%	72.5%	72.5%	72.5%	72.5%	80.3%	72.5%	72.5%	74.5%	74.5%	70.5%
Gsp229456 Native	70.5%	ID	98.0%	98.0%	98.0%	98.0%	98.0%	84.3%	98.0%	98.0%	96.0%	74.5%	72.5%
Gi197198	72.5%	98.0%	ID	100.0%	100.0%	100.0%	100.0%	86.2%	100.0%	100.0%	98.0%	76.4%	70.5%
Gi494	72.5%	98.0%	100.0%	ID	100.0%	100.0%	100.0%	86.2%	100.0%	100.0%	98.0%	76.4%	70.5%
Gi234179	72.5%	98.0%	100.0%	100.0%	ID	100.0%	100.0%	86.2%	100.0%	100.0%	98.0%	76.4%	70.5%
Gi240415	72.5%	98.0%	100.0%	100.0%	100.0%	ID	100.0%	86.2%	100.0%	100.0%	98.0%	76.4%	70.5%
Gi234328	72.5%	98.0%	100.0%	100.0%	100.0%	100.0%	ID	86.2%	100.0%	100.0%	98.0%	76.4%	70.5%
Gsp213198	80.3%	84.3%	86.2%	86.2%	86.2%	86.2%	86.2%	ID	86.2%	86.2%	84.3%	78.4%	80.3%
fascicula	72.5%	98.0%	100.0%	100.0%	100.0%	100.0%	100.0%	86.2%	ID	100.0%	98.0%	76.4%	70.5%
aggregatum	72.5%	98.0%	100.0%	100.0%	100.0%	100.0%	100.0%	86.2%	100.0%	ID	98.0%	76.4%	70.5%
Gsp240422	74.5%	96.0%	98.0%	98.0%	98.0%	98.0%	98.0%	84.3%	98.0%	98.0%	ID	78.4%	72.5%
cerebri	74.5%	74.5%	76.4%	76.4%	76.4%	76.4%	76.4%	78.4%	76.4%	76.4%	78.4%	ID	68.6%
G._rosea	70.5%	72.5%	70.5%	70.5%	70.5%	70.5%	70.5%	80.3%	70.5%	70.5%	72.5%	68.6%	ID

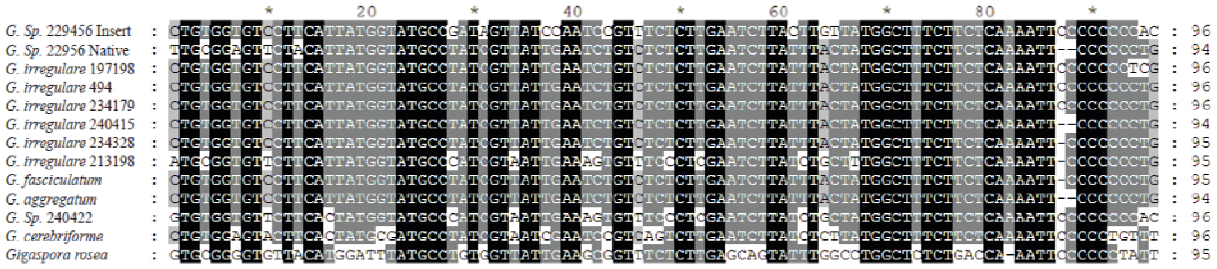


Figure S3.3. Multiple DNA sequence alignment of numerous AMF representatives of the *cox2* native C-terminals along with the G*lomus sp. 229456* putative foreign inserted C*- terminal.

Table S3.3. Sequence identity matrix of the *cox2* native C-terminals along with the *Glomus sp.* **229456 putative foreign inserted C*-terminal.**

Seq->	Gsp229456 Insert	Gsp229456 Native	Gi197198	Gi494	Gi234179	Gi240415	Gi234328	Gsp213198	fascicula	aggregatum	Gsp240422	cerebri	G._rosea
Gsp229456 Insert	ID	81.2%	87.5%	88.5%	88.5%	86.4%	87.5%	79.1%	87.5%	86.4%	83.3%	79.1%	63.5%
Gsp229456 Native	81.2%	ID	90.6%	92.7%	92.7%	94.6%	93.6%	85.2%	93.6%	94.6%	81.2%	78.1%	66.6%
Gi197198	87.5%	90.6%	ID	97.9%	97.9%	95.8%	96.8%	83.3%	96.8%	95.8%	84.3%	82.2%	65.6%
Gi494	88.5%	92.7%	97.9%	ID	100.0%	97.9%	98.9%	85.4%	98.9%	97.9%	85.4%	82.2%	66.6%
Gi234179	88.5%	92.7%	97.9%	100.0%	ID	97.9%	98.9%	85.4%	98.9%	97.9%	85.4%	82.2%	66.6%
Gi240415	86.4%	94.6%	95.8%	97.9%	97.9%	ID	98.9%	85.2%	98.9%	100.0%	83.3%	80.2%	64.5%
Gi234328	87.5%	93.6%	96.8%	98.9%	98.9%	98.9%	ID	86.3%	100.0%	98.9%	84.3%	81.2%	65.6%
Gsp213198	79.1%	85.2%	83.3%	85.4%	85.4%	85.2%	86.3%	ID	86.3%	85.2%	92.7%	73.9%	62.5%
fascicula	87.5%	93.6%	96.8%	98.9%	98.9%	98.9%	100.0%	86.3%	ID	98.9%	84.3%	81.2%	65.6%
aggregatum	86.4%	94.6%	95.8%	97.9%	97.9%	100.0%	98.9%	85.2%	98.9%	ID	83.3%	80.2%	64.5%
Gsp240422	83.3%	81.2%	84.3%	85.4%	85.4%	83.3%	84.3%	92.7%	84.3%	83.3%	ID	77.0%	61.4%
cerebri	79.1%	78.1%	82.2%	82.2%	82.2%	80.2%	81.2%	73.9%	81.2%	80.2%	77.0%	ID	59.3%
G._rosea	63.5%	66.6%	65.6%	66.6%	66.6%	64.5%	65.6%	62.5%	65.6%	64.5%	61.4%	59.3%	ID

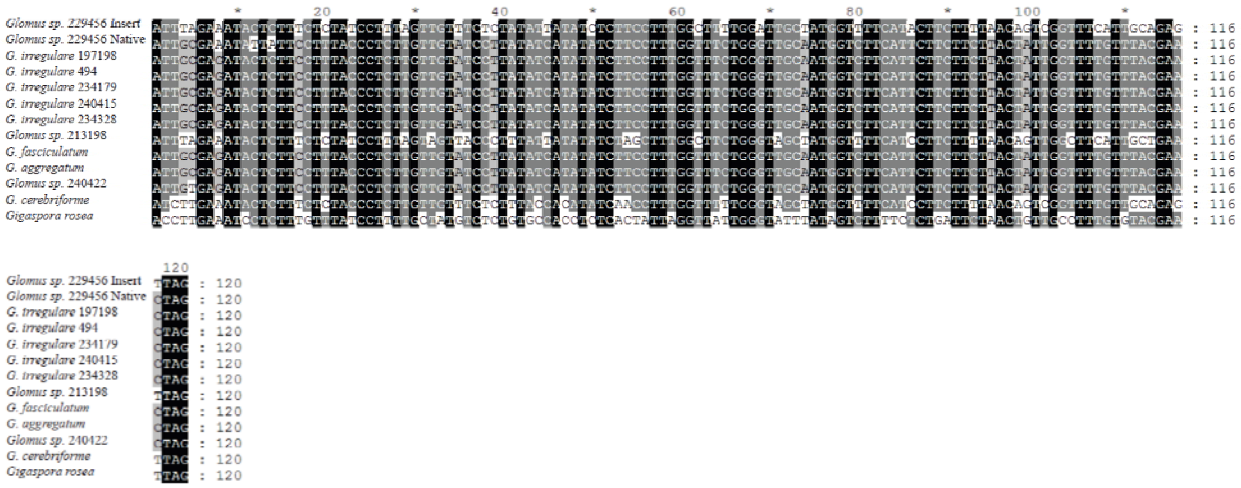


Figure S3.4. Multiple DNA sequence alignment of numerous AMF representatives of the *nad3* native C-terminals along with the *Glomus sp. 229456* putative foreign inserted C*- terminal

Table S3.4. Sequence identity matrix of the nad3 native C-terminals along with the Glomus sp. 229456 putative foreign inserted C*-terminal.

Seq->	Gsp229456 insert	Gsp229456 native	Gi197198	Gi494	Gi234179	Gi240415	Gi234328	Gsp213198	fascicula	aggregatum	Gsp240422	cerebri	G_rosea
Gsp229456 insert	ID	73%	74%	74%	74%	74%	74%	88%	74%	74%	74%	85%	61%
Gsp229456 native	73%	ID	98%	98%	98%	98%	98%	73%	98%	98%	97%	75%	63%
Gi197198	74%	98%	ID	100%	100%	100%	100%	73%	100%	100%	99%	76%	63%
Gi494	74%	98%	100%	ID	100%	100%	100%	73%	100%	100%	99%	76%	63%
Gi234179	74%	98%	100%	100%	ID	100%	100%	73%	100%	100%	99%	76%	63%
Gi240415	74%	98%	100%	100%	100%	ID	100%	73%	100%	100%	99%	76%	63%
Gi234328	74%	98%	100%	100%	100%	100%	ID	73%	100%	100%	99%	76%	63%
Gi213198	88%	73%	73%	73%	73%	73%	73%	ID	73%	73%	73%	83%	61%
fascicula	74%	98%	100%	100%	100%	100%	100%	73%	ID	100%	99%	76%	63%
aggregatum	74%	98%	100%	100%	100%	100%	100%	73%	100%	ID	99%	76%	63%
Gsp240422	74%	97%	99%	99%	99%	99%	99%	73%	99%	99%	ID	77%	64%
cerebri	85%	75%	76%	76%	76%	76%	76%	83%	76%	76%	77%	ID	68%
G_rosea	61%	63%	63%	63%	63%	63%	63%	61%	63%	63%	64%	68%	ID

Annex 3 : Supplementary Information (Chapter 4)

The mitochondrial genome of the glomeromycete *Rhizophagus* sp. DAOM 213198 reveals an unusual organization consisting of two circular chromosomes

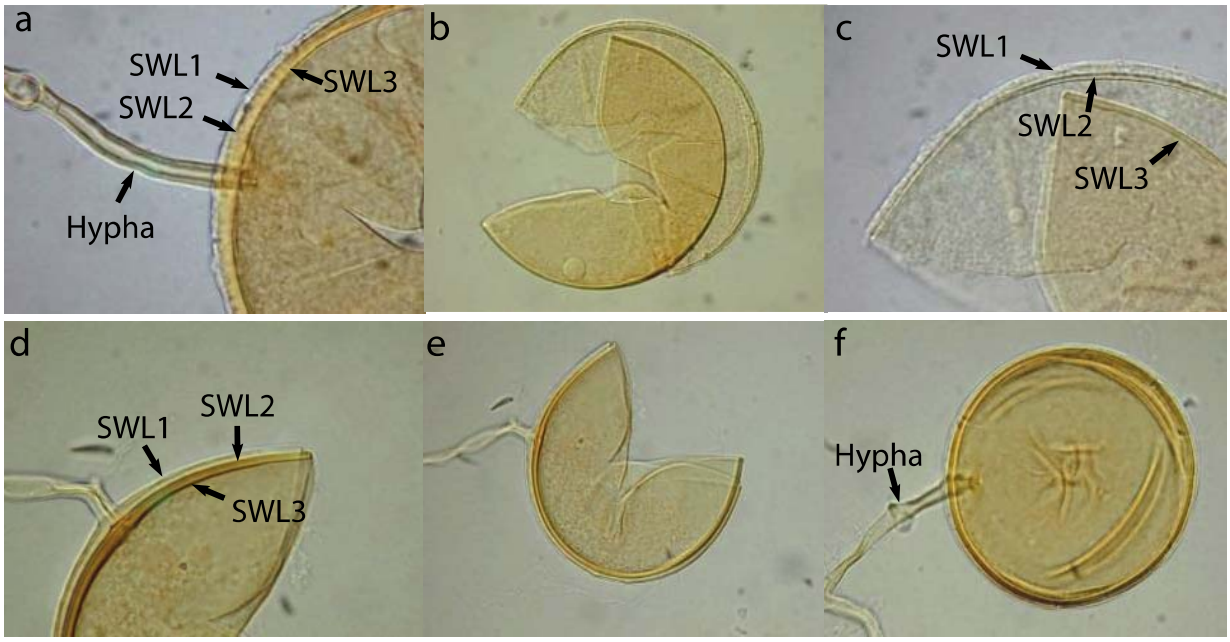


Fig. S4.1. Morphological description of spores of *Rhizophagus sp.* DAOM 213198 spores under *in vivo (a. b and c)* and *in vitro (e. f. g and h)* culture conditions.

The spore wall (SW) is composed of three layers in both *in vivo* and *in vitro* cultures (a and e). The outer layer SWL1 (b and c) was mucilaginous, evanescent, hyaline and approximately 1 µm thick. The second layer (SWL2) was rigid, smooth, hyaline and approximately 1.5 µm thick. The inner layer (SWL3) was laminate, smooth, pale yellow and approximately 1.5 µm thick. SWL1 and SWL2 were closely attached to each other forming a unique shell, easily detachable from the third layer (SWL3). The subtending hyphae (8.5–9.5 µm) was hyaline and straight and pore open.

These images are courtesy of Dr Yolande Dalpé (AAFC, Ottawa. ON).

2

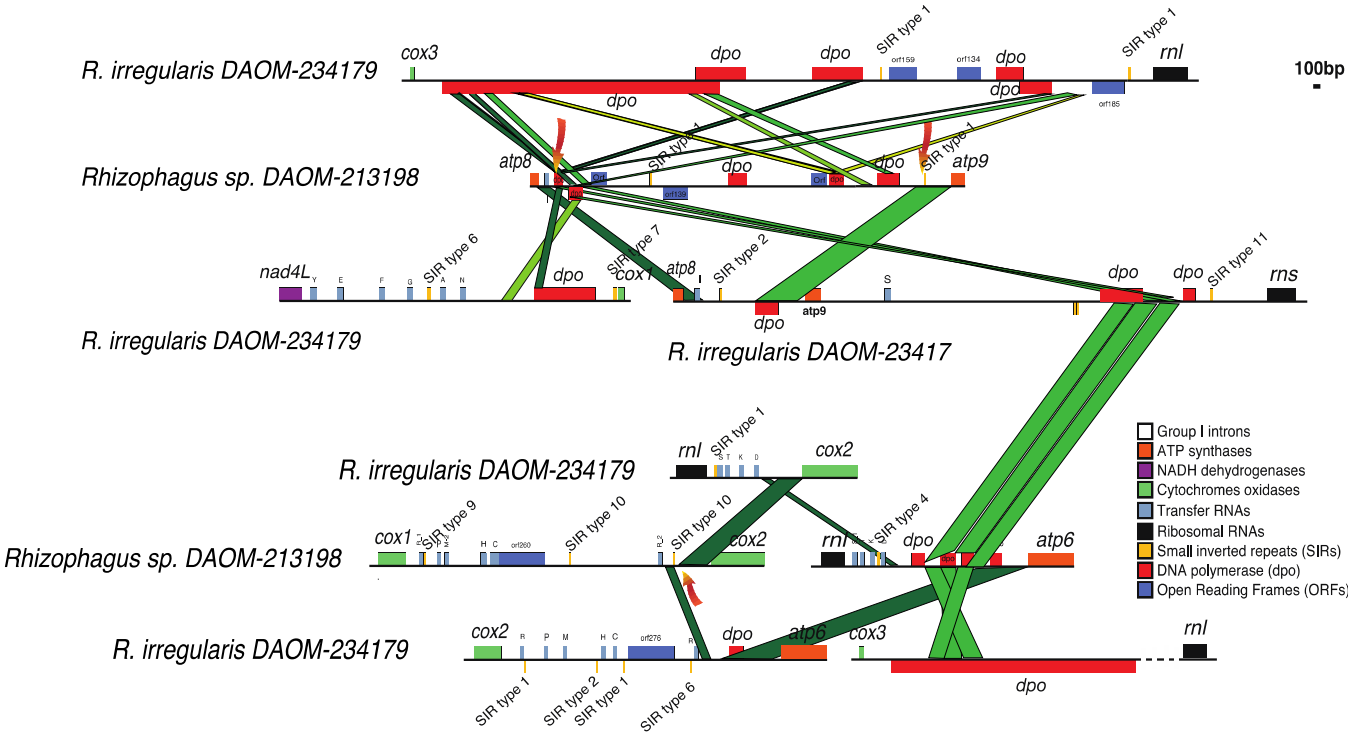


Fig. S4.2. mtDNA comparative analysis between *R. irregularis* DAOM 234179 and *Rhizophagus* sp. DAOM 213198 isolates.

The linear-mapping of mtDNAs containing the newly formed and reshuffled intergenic regions of *rnl-atp6* and *cox1-cox2* and *atp9-atp8* in *Rhizophagus* sp. DAOM 213198 in comparison to its relative *R. irregularis* DAOM 234179. Nucleotide identity comparison with tBLASTx, between the homologous regions is indicated by the projections. Dark green projections represent sequence identity higher than 90% while light green projections represent sequence identity ranging between 50 and 90%.

3

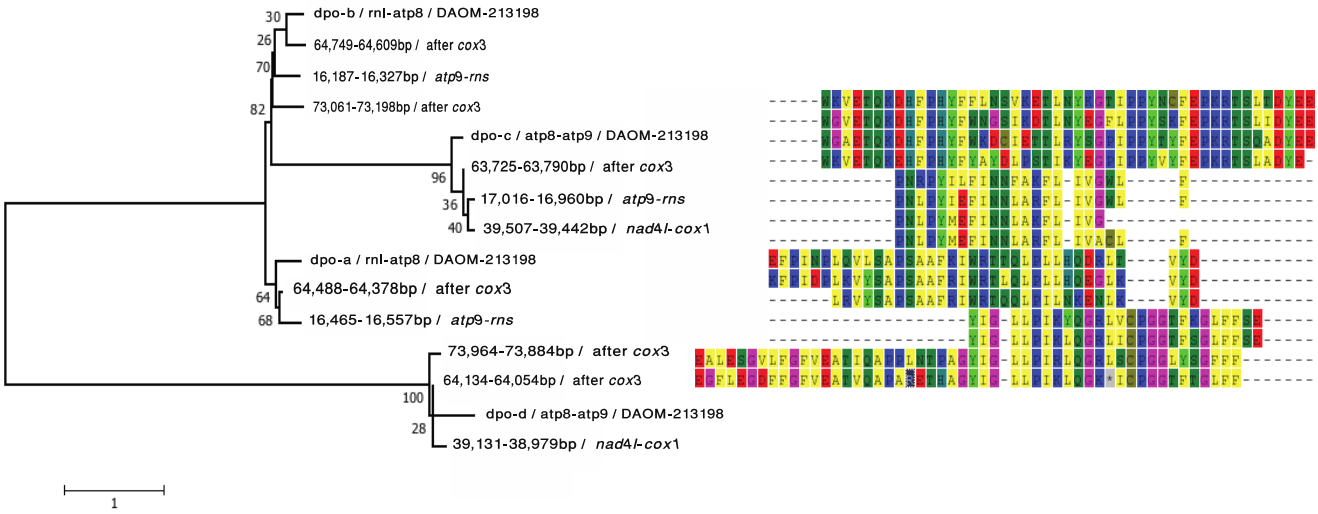


Fig. S4.3. Alignment of potentially conserved *dpo*-like translated amino acid sequences from *R. irregularis* DAOM 234179 and *Rhizophagus* sp. DAOM 213198.

Distance sequence analysis of *dpo*-like fragments of two closely related species of *R. irregularis* DAOM 234179 and *Rhizophagus* sp. DAOM 213198 reveals that there are some conserved domains in *dpo*s positioned in different intergenic regions.

4

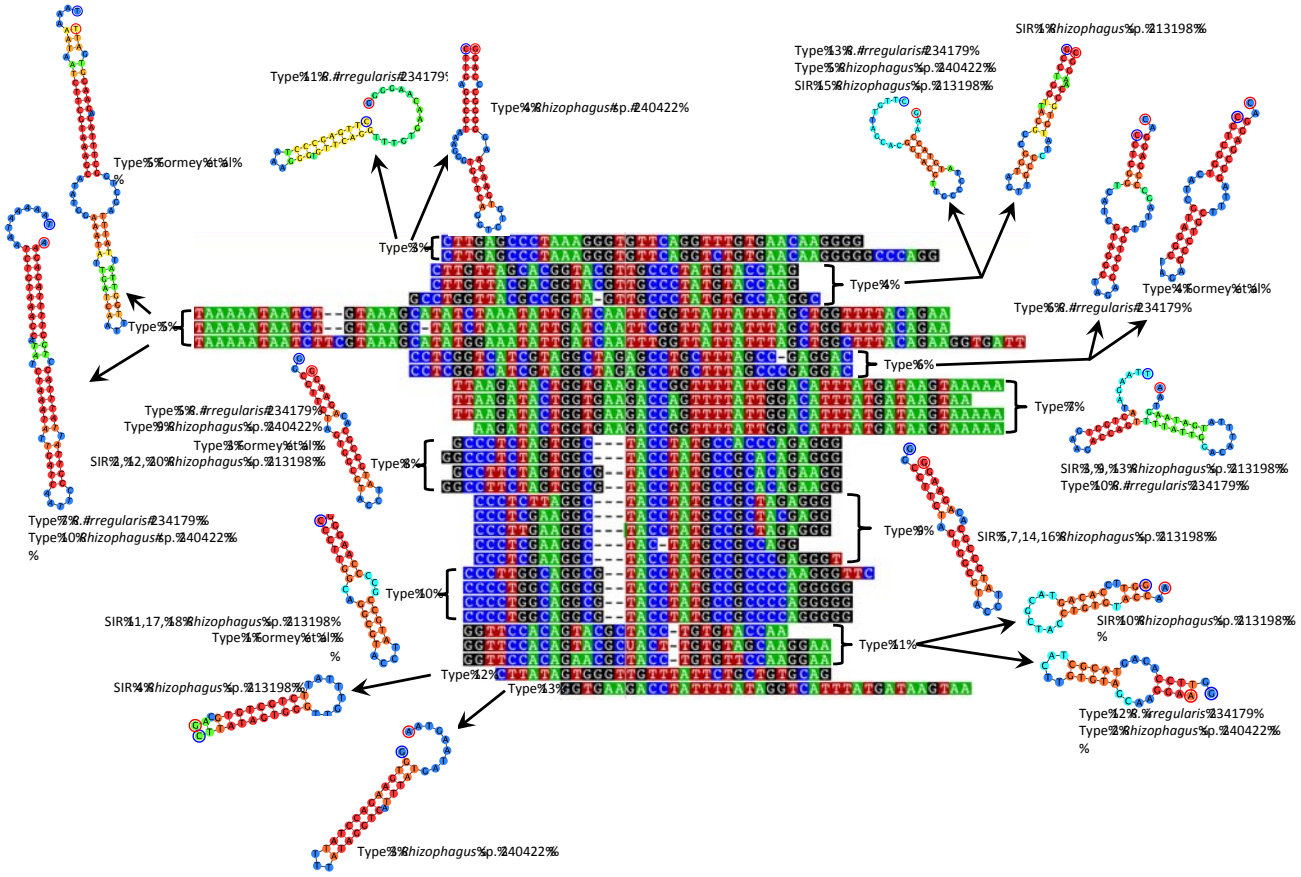


Fig. S4.4. Multiple sequence alignment of less frequent short inverted repeats.

The previously identified SIRs subtypes found in *R. irregulare* DAOM 234179, *Rhizophagus* sp. DAOM 240422 and 213198 (Formey *et al.* 2012) were aligned and classified. Types 3 to 13 represent less frequently found SIRs subtypes in DAOM 213198. Secondary structures were predicted based on the energy model of Mathews *et al.* (2004) and Andronescu *et al.* (2007). The minimum free energy (MFE) structure of hairpins is colored by base-pairing probabilities (red: high; green: mid; blue: low). Blue and red circles around nucleotides represent the beginning and the end of molecule, respectively.

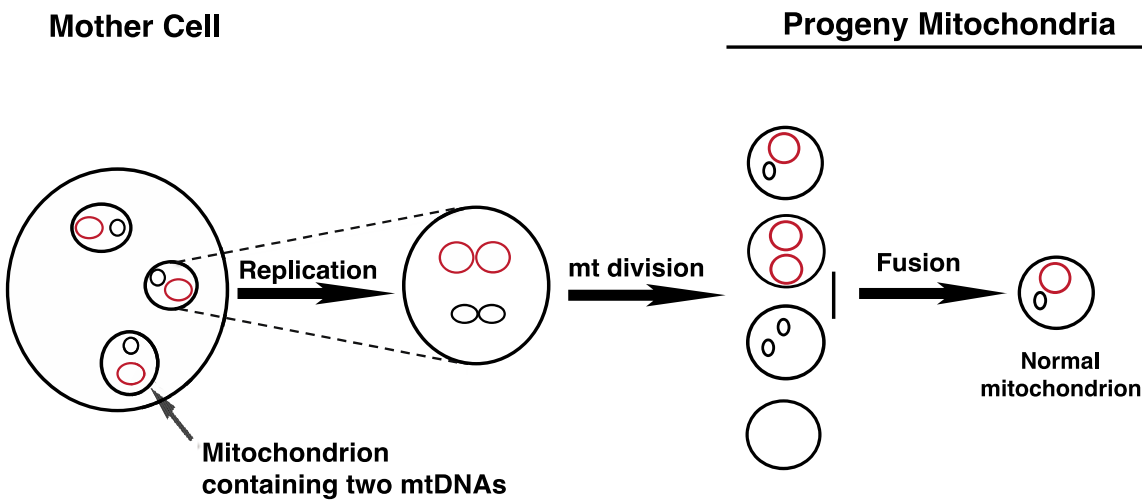


Fig. S4.5. Hypothetical pathway of mtDNA inheritance and dynamics in *Rhizophagus* sp. DAOM 213198.

Mitochondrial division could result in the formation of mitochondria containing either mtDNAs or one of each mtDNAs. Fusion dynamic of mitochondria also provide the opportunity for the ones harboring incomplete mtNA to fuse and complete set of mitochondrial genes.

Table S4.1. Absolute quantitative real-time PCR assays performed on DAOM 213198 and DAOM 197198 using *cox1* and *rnl* for the large and small mtDNAs, respectively. PCR efficiencies were 98.33% and 99.86% for *rnl* and *cox1*, respectively.

		rnl quantity	*cox1* quantity	ratio *cox1/rnl*
293198	Dilution 1	27 113	27 852	1.0
	Dilution 2	2 468	2 668	1.1
	Dilution 3	236	291	1.2
	Dilution 4	30	23	0.8
197198	Dilution 1	3 224	2 459	0.8
	Dilution 2	295	231	0.8
	Dilution 3	26	20	0.8

Annex 4 : Supplementary Information (Chapter 5)

Detection of a transient mitochondrial DNA heteroplasmy in the progeny of crossed genetically divergent isolates of arbuscular mycorrhizal fungi

Figures S5.1. And S5.2. monosporal culture lines of *Rhizophagus irregularis*

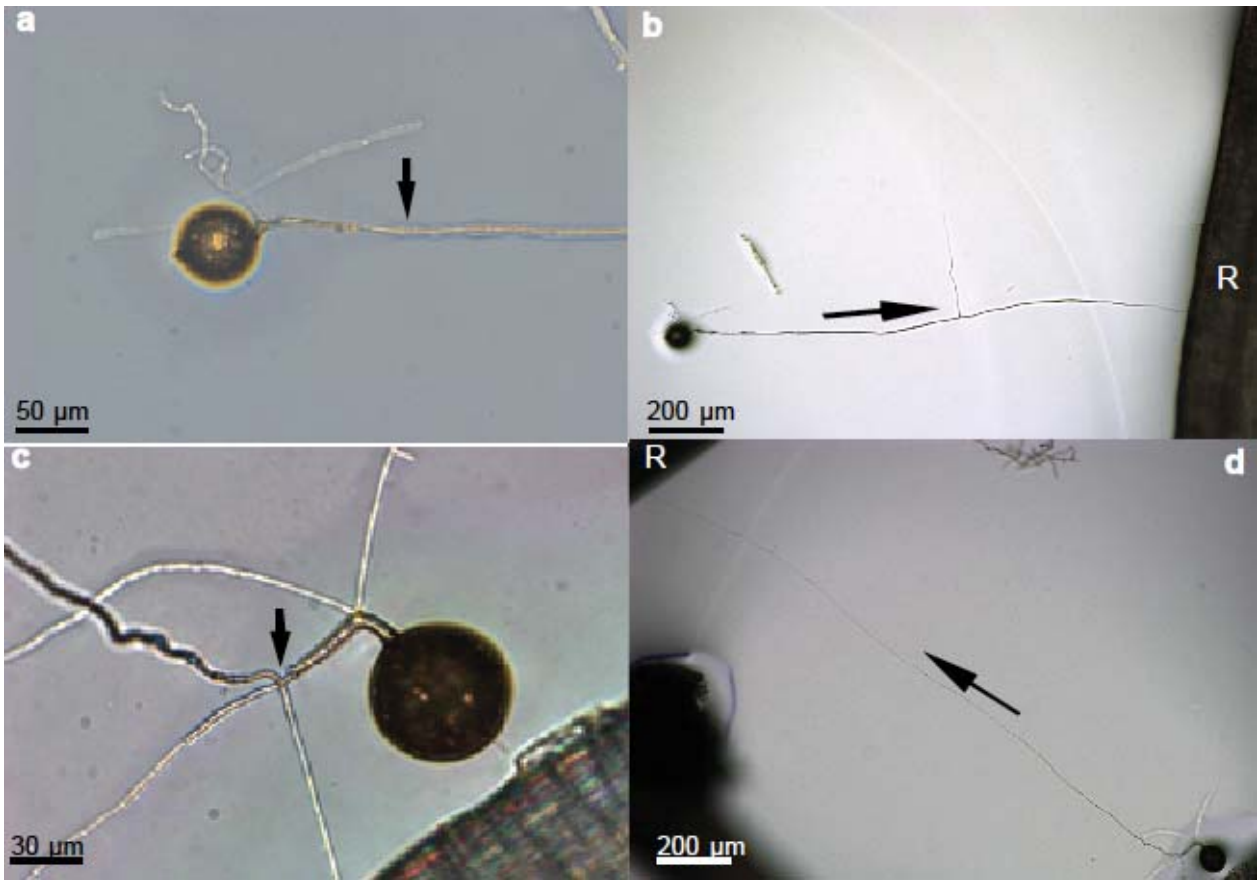


Figure S5.1. Spore germination of *R. irregularis* through the subtending hyphae.
Black arrowhead and rhizotropism (i.e. growth oriented towards the roots (R) (black arrow) related to attempts to achieve monosporal culture lines from progenies originating from crossed parents. Panels a and b: spore progenies originating from combination DAOM-197198/DAOM-234328; c and d: spore progenies from combination DAOM-234328/DAOM-240415.

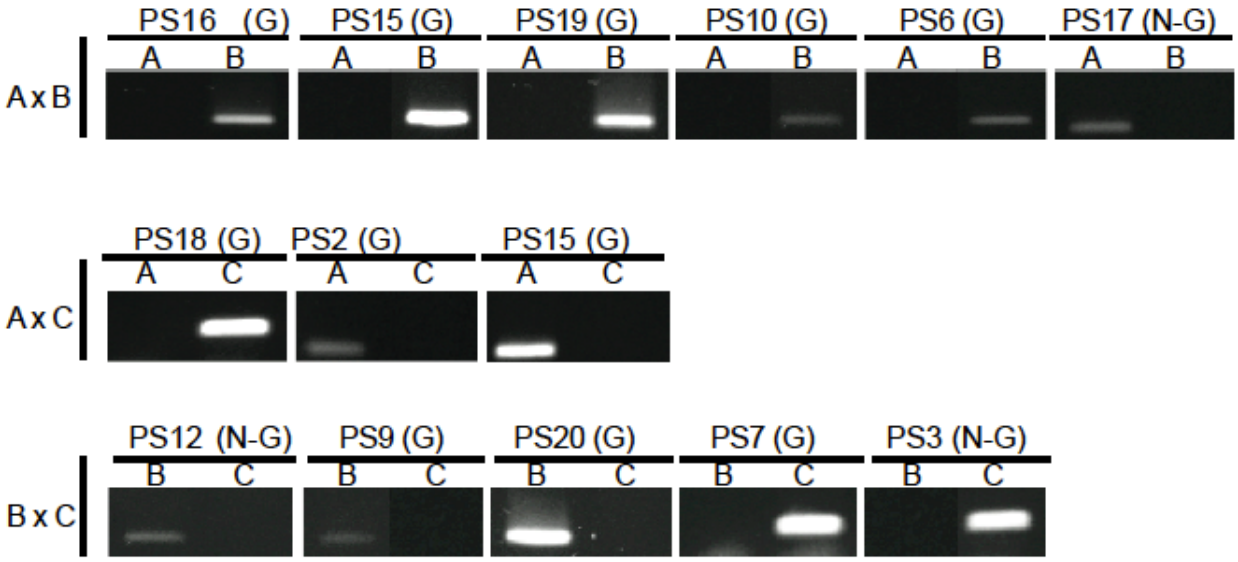


Figure S5.2. Polymerase chain reaction banding patterns of the progenies coming from different parental combinations (AB, AC, BC).

Each progeny spore from the monosporal cultures lines are indicated by (PS), which has been tested by both parental specific primers. *R. irregularis* isolates DAOM-197191 (A) DAOM-240415 (B) and DAOM-234328 (C). mtDNA marker sizes are 120 bp, 147 bp and 209 bp corresponding to DAOM-197198, DAOM-240415 and DAOM-234328, respectively.

Table S5.1. Number and percentage of anastomosis[1] between germlings from spore clusters either from the same or different isolates of *R. irregularis*.

Type of interaction Within the same isolate	Number of contacts	Perfect fusion (Self fusion)*(%)	No interaction	Pre-fusion incompatibility or hyphal avoidance (%)	Post-fusion incompatibility (%)
A-A	205	123 (60)	82	0	0
B-B	175	68 (38)	107	0	0
C-C	243	135 (55)	108	0	0
Between isolates	Number of contacts	Non-self fusion* (%)	No interaction	Non-self pre-fusion incompatibility or hyphal avoidance (%)	Non-self post-fusion incompatibility (%)
A-C	148	0	145	0	3 (2)
A-B	134	0	130	3 (2,3)	1 (0,7)
B-C	157	2* (1,2)	151	2 (1,3)	2 (1,3)

A- DAOM 197198, B- DAOM 240415, C- DAOM 234328

* Fusions were followed for three weeks and no septa were formed between the connected hyphae.

[1]Anastomoses within the same mycelium in crossing treatment were not recorded.

Table S5.2. Total number of spores[1] of *R. irregularis* (full lipids, empty and aborted like-structures) produced in the interaction zone and at both sides in the crossing experiments and controls.

Combination	**A[2]**	**Int-Zone**	**B**	**B**	**Int-Zone**	**C**	**A**	**Int-Zone**	**C**
Full lipids	1329 ± 94	21± 8	1149 ± 89	1019±141	21±4	1638±66	1497±329	14±7	1317±267
Empty	27± 4	1± 0,5	17± 5	15±4	1±0,5	27±6	34±8	3±3	30±11
Aborted	20±4	2±1	13±6	19±12	0	15±8	14±3	2±1	23±3
Total	1376±90	25±10	1179±94	1054±152	22±4	1682±79	1546±322	19±11	1371±279
Control	**A**	**Int-Zone**	**A**	**B**	**Int-Zone**	**B**	**C**	**Int-Zone**	**C**
Full lipids	1249±317	32±15	1582±125	1807±345	46±24	1899±212	1949 ± 84	45 ± 11	1584 ± 461
Empty	22±12	1±1	14±10	24±11	2±1	36±10	26 ± 4	0	7 ± 4
Aborted	16±9	5±1	9±4	25±9	5±2	28±8	17 ± 1	2 ± 2	23 ± 9
Total	1288±338	39±14	1604±139	1857±359	52 ± 27	1964±220	1988 ± 108	47 ± 12	1605 ± 473

A- DAOM 197198, B- DAOM 240415, C- DAOM 234328

[1]No significance differences were found within the treatments and controls (P≥0.05)

[2] Mean ± Standard Error

Table S5.3. Number and percentage of germinated and non-germinated spores of *R. irregularis*, which were tested to produce monosporal culture lines.

Combination	Number of offprings tested	Germinated offprings (%)	Progeny which contacted the root after germination	Offprings showing heteroplasmy (%)
A-C	20	40	5	0
A-B	20	50	10	0
B-C	20	60	20	0

A- DAOM 197198, B- DAOM 240415, C- DAOM 234328