Clues of Sexual Reproduction in the 'Ancient Asexual' Fungal Lineage: the Arbuscular Mycorrhizal Fungi

Rohan Riley

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
University of Ottawa
In partial fulfillment of the requirements for the
M.Sc. degree in the
Ottawa-Carleton Institute of Biology

Thèse soumise à la
Faculté des études supérieures et postdoctorales
Université d'Ottawa
En vue de l'obtention de la maîtrise ès science
L'Institut de biologie d'Ottawa-Carleton

Abstract

Arbuscular mycorrhizal fungi (AMF) represent an ancient and critical symbiotic partner with the majority of land plants, understood to promote ecosystem productivity and biodiversity and are also important to ecologically sound land management practices. The fungus is thought to be over 400 million years old, and due to a lack of an observable sexual cycle, has been placed into a select group of eukaryotes called 'ancient asexuals', which seemingly defy evolutionary theory by persisting for an extended period of time in the absence of sexual reproduction. Recently however, molecular evidence has accumulated which may suggest AMF harbour a cryptic sexual cycle. In the first chapter of this thesis, entitled "Searching for clues of sexual reproduction in the genomes of arbuscular mycorrhizal fungi", I review evidence supporting this notion of a cryptic sexual cycle in AMF which includes: the presence of recombinational events, meiosis-specific genes as well as mating-specific transcription factors called *SexM* and *SexP* of the MATA_HMG protein family which are otherwise found only in the genomes of sexual fungi. In the second chapter, I present the main research of my MSc work where I used bioinformatic, population genetic, molecular and experimental approaches to build on this existing evidence of sexuality in AMF. These findings include the presence of a dramatically expanded family of *MAT*-HMG genes which are present in several isolates of the *Rhizophagus irregularis* and also harbour significant allele variation amongst these isolates, some of which resembles variation expected at *MAT*-genes in other fungi. Q-RT-PCR procedures revealed that at least some of these genes tend to increase in expression during crosses of *R. irregularis* isolates. We also uncovered the presence of a unique genomic region where at least three of these genes are located in tandem. Finally, several tests of recombination support the presence of intraisolate as well as interisolate recombination events occurring between these *MAT*-HMG genes.

Resumé

Les champignons mycorrhiziens à arbuscules (CMA) représentent des anciens partenaires symbiotiques de la majeure partie des plantes terrestres, ce des organismes qui promeuvent la productivité et la biodiversité de nombreux écosystèmes terrestres. Ces champignons sont présents depuis 400 millions d’années, mais malgré leur ancienneté une sexualité n’a jamais été observée chez ces organismes. Pour cette raison, ces champignons sont considérés comme faisant partie d’un groupe appelé " anciens asexués ". Néanomoins, de nombreuses séquences d’ADN de CMA ont récemment, révélé la présence de gènes liés à la sexualité chez ces organismes, et cette découverte suggère la présence d’un sexualité cachée chez les CMA. Le premier chapitre de ma thèse est intitulé "Recherche d’indices de sexualité chez les champignons mycorrhiziens à arbuscules", et dans celui-ci je passé en revue les récentes découvertes qui supportent la présence d’un cycle sexué chez les CMA. Dans le deuxième chapitre, je présente la recherché principale effectuée en laboratoire, ou des approches de bioinformatique, génétique des populations, moléculaires et expérimentales ont été utilisées pour démontrer la présence d’un cycle sexué chez les CMA. Ces découvertes incluent la découverte d’un énorme nombre de de gènes *MAT*-HMG dans le génome des ces champignons, ce qui est intéressant car ces gènes jouent généralement un rôle important dans la sexualité d’autres champignons. J’ai aussi démontré que ces gènes sont présents sous différentes formes alléliques entre individus d’une même population, comme chez d’autres champignons sexués, et que certains de ces gènes sont fortement surexprimés lors de croisement sous conditions expérimentales. Mes analyses génomiques ont aussi permis la découverte d’une région unique dans le génome des CMA, composée de trois gènes MAT-HMG présents en tandem. Nos analyses ont aussi révélé la présence de recombinaison entre souches d’une espèce de CMA.

Table of Contents

Acknowledgments..*1*

Introduction...*2*

ArbuscularMycorrhizal Fungi *2*

The AMF Life Cycle *3*

AMF Ecology..*4*

AMF classification and identification..*7*

Cellular and molecular biology of AMF...*8*

Chapter 1: Searching for clues of sexual reproduction in the genomes of arbuscular mycorrhizal fungi *13*

Abstract *13*

Introduction..*14*

Known signatures of sex in the AMF genome..*16*

Further exploring cryptic sex in AMF...*20*

Chapter 2: Identification of an expanded family of fungal *MAT*-HMG genes in the arbuscular mycorrhizal fungus *Rhizophagus irregularis*...*27*

Abstract *28*

Introduction..*28*

Materials and Methods..*32*

Results...*49*

Discussion...*78*

Concluding remarks...*83*

Conclusions...*84*

References...*90*

Acknowledgments

I would like to thank my supervisor, Dr. Nicolas Corradi for his support and motivation throughout this project. I would also like to extend this thanks to my committee members, Dr. Linda Bonen, Dr. Myron Smith, and Dr. Alexandre Poulain for their helpful advice throughout the course of this project. Special thanks to Dr. Stéphane Aris-Brosou for teaching me some of the essentials of phylogenetics. Thanks to all current and former lab members for their support; Mohammed Selman, Adrian Pelin, Steve Ndikumana, Stephanie Pichette, Stephan Amyotte, Carolin Sachgau and especially Phillip Charron for his work on this project. Finally, a special thanks to the collaborators for their instrumental role in this project: Dr. Francis Martin for providing AMF sequence data, Dr. Yolande Dalpé for providing us with DNA from AMF isolates originating all over the world, and Dr. Alexander Idnurm for carrying out inverse PCR.

Introduction

A Symbiosis is defined as an intimate association between two species (Smith *et al.* 1997) and has played a major role in shaping life on earth. Lichens, for example, have played a fundamental role in soil formation over long periods of time and currently are the result of a symbiotic association between cyanobacteria and fungi (Banfield *et al.* 1999). A hallmark of eukaryotic cells are the presence of mitochondria and chloroplasts which have evolved from bacterial endosymbionts. These are responsible for cellular respiration and photosynthesis, respectively, and have been major factors for the success of eukaryotic cells (Margulis 1981).The Kingdom Fungi has many examples of symbiotic associations where some are beneficial and others detrimental to one of the partners, which is a range commonly referred to as the “symbiotic continuum” (Johnson *et al.* 1997; Rodriguez and Redman 2008).

The Mycorrhizal symbioses, derived from the Greek words 'mykes' meaning fungus, and 'rhiza' meaning root, specify plant-fungal associations with 'dual organs of adsorption formed when symbiotic fungi inhabit healthy tissues of most terrestrial plants' (Johnson, 1997).This association can be both parasitic or mutualistic, and can be found in a variety of environments ranging from tropical forest to temperate grassland. Mycorrhiza involve diverse fungal taxa within the fungal phyla; namely the Basidiomycota, Ascomycota and Glomeromycota.(Kottke and Nebel 2005; Schussler *et al.* 2001).

Arbuscular Mycorrhizal Fungi

Arbuscular mycorrhizal fungi (AMF) are an ancient and geographically widespread group of plant symbionts, belonging to the phylum Glomeromycota, forming what is known as the Arbuscular Mycorrhizal Symbiosis with the majority of land plants and many economically important crops (Bonfante and Selosse 2010; Corradi and Charest 2011; Sanders 2002; Sanders 2003). This symbiosis appears to be ancient, as fossils of AMF dating back approximately 460 million-years ago contain morphological fungal structures that are identical to those of extant counterparts (Humphreys *et al.* 2010; Redecker *et al.* 2000). In some cases, fossilized remains depict the AMF forming an association with early land plants that resemble modern-day bryophytes, such as the liverwort (Remy *et al.* 1994).This Fossil evidence, together with molecular dating (Simon et al., 2001) and the widespread geographical distribution of the symbiosis, supports the current hypothesis that the AM symbiosis was crucial to the colonization of land by plants (Pirozynski,1966). In addition to being symbionts of vascular plants and bryophytes, an early diverging AMF is associated with cyanobacteria - *Geosiphon pyriforme.* This association suggests that the early evolution of AMF could have begun with a symbiotic association between a cyanobacteria or algae before establishing themselves as major symbionts of land plants (Gehrig *et al.* 1996).

The AMF life Cycle

The AMF symbiosis is thought to be initiated by a cross-talk of diffusible signals between the fungal and plant partners (Harrison 2005). The life cycle of AMF begins as a resting spore in the soil, followed by germination of hyphal filaments - the fungal equivalent of cells- in search of a host root. Hyphal growth in the absence of the plant may occur for one to two weeks where an AMF may terminate growth to store energy before attempting a second germination (Bago *et al.* 2000). Upon root contact, the fungus forms an 'infection' apparatus called an appressorium on the root surface, which initiates intraradical growth of the fungus where the symbiotically active phase of the lifecycle occurs. This is marked by the formation of arbuscules, the primary site of nutrient and carbon exchange between the plant and fungi along the vascular tissue of the plant and lipid-filled vesicles which function as energy storage. Growth within the root is followed by extraradical growth out of the root and into the soil environment, where the fungus scavenges for nutrients and water, forming spores and colonizing additional root tissue (Friese and Allen 1991; Rillig 2004).

AMF Ecology

Effects on the individual plant

As an obligate plant symbiont, growth of the AMF is driven by photosynthetically derived plant carbohydrates which are exchanged for water and mineral nutrients, including calcium, potassium, zinc and most notably phosphorus and nitrogen - two nutrients that can be very limiting for plant growth (Friese and Allen 1991).These, scavenged by the fungus, result in a substantial expansion of the plant rhizosphere beyond the root`s zone of nutrient depletion where AMF hyphae can occupy an estimated 111m cm^{-3} of soil (Miller *et al.* 1995). The net result usually favours the plant, leading to increases in plant growth, tolerance to pathogens, pests and environmental stresses such as drought and salinity (Smith *et al.* 1997). In fact, these beneficial effects of AMF have attracted great interest in the application of AMF in organic agriculture, such that several companies now produce commercial inocula of a single isolate of *Rhizophaus irregularis*.

Effects of AM symbiosis on soil health and carbon storage

AMF are an important contributor to the formation and maintenance of soil health and account for a significant sink of carbon. AMF mycelia produce glomalin, a substance secreted by the hyphae that binds soil particles together to form stable macroaggregates and represents approximately 5% of total soil carbon (Zhu and Miller 2003). These macroaggregates form the basic building blocks of soil structure; they hold in soil organic matter which affects the overall nutrient and water holding abilities of the soil (Jeffries *et al.* 2003) and serve as microhabitats for many soil organisms, including bacteria, nematodes, and arthropods. Furthermore, carbon delivered to the soil by AMF hyphae increases the activity of other soil biota and, curiously, this effect appears to be selective to soil biota antagonistic to soil-born plant pathogens (Jeffries *et al.* 2003)

Colonization of AMF results in a 4-20% transfer of the photo-assimilates to the plants mychorrhizal root system, and carbon stored by the hyphae accounts for an estimated 50 to 900 kg ha^{-1} of soil carbon content (Zhu and Miller 2003). At the extreme, one estimate found that the AMF symbiosis is responsible for the flow of 5 billion tons of carbon annually on a global scale (Bago *et al.* 2000).

Effects of AMF biodiversity on plant biodiversity

Belowground AMF biodiversity has been suggested to have major effects on aboveground plant biodiversity and productivity. AMF taxa have been shown to have host preferences to different plant species (Vandenkoornhuyse *et al.* 2003) and functional compatibilities, such that the amount of phosphorous uptake provided by a single AMF species depended on plant identity (Ravnskov and Jakobsen 1995). Furthermore, differing AMF taxa have been found to provide different benefits to the host in terms of pathogen resistance (Newsham *et al.* 1995) and phosphorus uptake (Jansa *et al.* 2005; Munkvold *et al.* 2004)

AMF have also been shown to reduce interspecific competition between coexisting plant species. For example, Wagg et al. (2011) demonstrated that competition between *Trifolium pratense* and *Lolium multiflorum* (a legume and a grass) grown in the same pot differed in the presence or absence of AMF. Specifically, in the absence of AMF the competition with *L. multiflorum* caused a large growth reduction in *T. pratense* relative to *T. pratense* growing in the absence of *L. multiflorum*, while the presence of AMF resulted in an increased growth of *T. pratense* under competition with *L. multiflorum* compared to the growth of *T. pratense* in the same condition in the absence of AMF. Moreover, under competition of *L. multiflorum*, the increase in growth of *T.pratense* changed depending on the AMF species inoculated (Wagg *et al.* 2011). Thus the presence of AMF facilitates the coexistence of plant species, potentially resulting in a more diverse community of plant species to coexist.

Greater diversity of AMF may improve ecosystem productivity through a mechanism called functional complementarity, where different AMF species - which differ in the various host benefits they provide - could sum up to a greater plant benefit than any one individual AMF can provide alone. Van der Heijden *et al.* (1998) observed a greater plant community phosphorus uptake with an increase in AMF species diversity. They hypothesized this occurred through a greater overall hyphal length provided by different AMF species which enabled the plants to scavenge the surrounding soil for phosphorus more efficiently. Indeed, Smith *et al.* (2000) demonstrated that two AMF species differed in the zone around the root from which they extracted phosphorus. Moreover, the same variation in AMF scavenging strategy was shown to occur between isolates of a single AMF species (Munkvold *et al.* 2004). These authors hypothesized that the observed variations in soil scavenging strategies for phosphorus by different AMF individuals might provide an overall increase in above ground productivity when greater number of AMF taxa or isolates are present (Munkvold *et al.* 2004; Smith *et al.* 2000).

AMF diversity influences competitive interactions of above ground plants, and provides a variation in specific host benefits, likely leading to functional complementation when greater diversity of AMF are present. Overall, this translates to increases in aboveground plant biodiversity which also has been shown to positively influence ecosystem productivity (Tilman *et al.* 1996; van der Heijden *et al.* 1998). Thus AMF are a keystone species in terrestrial ecosystems.

AMF classificationand identification

Initially AMF were grouped into the class of fungi *Glomeromycetes* within the paraphyletic phyla Zygomycota, based purely on morphological and developmental similarities to other members of the group (Cavalier-Smith 1998). A more comprehensive analysis using the small subunit of the ribosomal RNA (SSU rDNA), later resulted in AMF being included in their own new phylum, the Glomeromycota (Schussler *et al.* 2001). For some time, Glomeromycota have been thought to be closely related to higher fungal phyla, specifically the Ascomycota and the Basidiomycota based on these ribosomal genes (Schussler *et al.* 2001). Current phylogenies constructed using concatenations of many nuclear and mitochondrial genes, have suggested their closer association with more basal members of the fungal Kingdom, namely the Mortierellales (Zygomycota) (Corradi and Sanders 2006; Halary *et al.* 2011; Lee *et al.* 2009; Liu *et al.* 2009; Liu *et al.* 2006; Nadimi *et al.* 2012; Pelin *et al.* 2012).

To date, AMF include a total of 280 species in 19 genera. The highly reduced morphological diversity of AMF has resulted in taxonomical difficulties but molecular methods are beginning to circumvent these issues. Current methods for species identification combine morphological studies with genetic markers, the latter being based on a combination of the Small Subunit (SSU) of the nuclear DNA, the mitochondrial large subunit (mtLSU) genes and nuclear internal transcribed spacers (ITS).These markers have now proven useful for AMF identification down to the species level (Kruger *et al.* 2012). Undoubtedly next generation sequencing technologies will further facilitate our understanding of AMF diversity and community composition.

Cellular and molecular biology of AMF

The molecular biology of AMF has been notoriously difficult to study, resulting in an overall long-term poor understanding of their genetics and evolution. For a long time, the major difficulty in studying AMF genetics was to obtain “contaminate-free” fungal material from different species due to the inability to culture AMF in the absence of host roots; this has been partially circumvented with the development of *Agrobacterium tumerficans* transformed roots that can form a symbiotic relationship with AMF under axenic conditions without (i.e. *in-vitro*)without the need for photosynthesis.

Ploidy and genome size estimates

Studies of ploidy levels in AMF have suggested that at least one species (*R. irregularis)* is haploid (Hijri and Sanders 2004; Martin *et al.* 2008). However, large ranges in genome size estimates among several species of AMF have led to speculations that some AMF may be highly polyploid (Pawlowska and Taylor 2004; Sanders and Croll 2010). Current estimates of genome size for AMF vary considerably. For instance, one single isolate of *R. irregularis* has been proposed to have a genome size of either 15Mb , or 150 Mb depending on the authors (Hijri and Sanders 2004; Sędzielewska *et al.* 2011). On the large end of the genome size spectrum, the AMF *Gigaspora Margerita* has a genome estimated at 740Mb (Hosny *et al.* 1998).

The cellular and molecular features of AMF

Several cellular and molecular features of AMF put them at odds with views of conventional eukaryotes. AMF cells are represented by ceonocytic, filamentous hyphae that are perpetually multinucleated, which means several nuclei of one individual along with other cytoplasmic elements can freely flow through one common cytoplasm. AMF are also capable of producing anastamoses (Croll *et al.* 2009; Giovannetti *et al.* 1999), a mechanism where hyphae of the same or different individuals fuse to form a cytoplasmic bridge (Figure 1). This is thought to allow resource sharing between hyphal networks of different individuals and has been proposed as a mechanism for the exchange of nuclei between individuals. Interestingly, extensive genetic polymorphisms have been observed within a single species of AMF, raising the hypothesis that the coexisting nuclei in a mycelial network may be present in anheterokaryotic state (i.e. genetically different nuclei co-existing in one cytoplasm). Distribution of these polymorphisms has been suggested to result from other mechanisms (i.e. polyploidy), and this particular subject has been extensively debated (Hijri and Sanders 2005; Kuhn *et al.* 2001; Lanfranco *et al.* 1999; LloydMacgilp *et al.* 1996; Pawlowska and Taylor 2004; Rodriguez *et al.* 2004; Stukenbrock and Rosendahl 2005).

Figure 1. Micrograph of two anastamosing hyphae of *Funneliformis mossae.* Scale bar is 10um. (Giovannetti *et al.* 1999).

Functional investigations are made difficult by the unavailability of an appropriate transformation system for AMF possibly due to their atypical filamentous structure. Successful reports have amounted to transient phenotypes in two species of AMF (Harrier 2001; Sanders and Croll 2010;Tisserant *et al.* 2012), but complete transformation has yet to be successful. This has limited some functional investigations of genes in AMF to transformation experiments involving mutants of other fungal species lacking the homologue of interest (Tollot *et al.* 2009). A recently developed method called Host Induced Gene Silencing (HIGS) (Nowara *et al.* 2010) which relies on an expressed RNAi construct in the host root to silence the fungal partners transcripts has yielded some positive results (Helber *et al.* 2011). It remains to be seen if this method will be broadly applicable to other AMF transcripts.

Ancient asexuals? Or maybe not.

It is clear that individual AMF are capable of exerting unique effects on their environments (e.g. by differential effects on plant growth). The mechanisms that produce this genetic diversity and ultimately lead to different phenotypes in AMF populations are largely unknown. In eukaryotes, sexual reproduction is initiated by the fusion of two haploid genomes to form a diploid zygote, and ends with a reduction division back to haploidy following meiosis and recombination. This mechanism is widespread in the eukaryotes and is currently thought to be at least one billion years old (Perrin 2012). Curiously, AMF have long been suggested to have evolved for over 400 million years in the absence of sexual reproduction which would place these fungi into the highly select group of eukaryotes called the 'ancient asexuals' (Judson and Normark 1996). Nevertheless, evidence now starts to emerge that the AMF may be able to undergo a cryptic sexual cycle, and these are catalogued across the first chapter of this thesis, entitled: “Searching for clues of sexual reproduction in the genomes of arbuscular mycorrhizal fungi”.

In the second chapter, entitled ``The identification and analysis of an expanded family of mating-specific transcription factors in the AMF *Rhizophagusirregularis*``, I present the experimental portion of this thesis which builds on this prior evidence of sex in AMF. Specifically, the research presented in this chapter aimed to identify a mating-type specifying locus within the AMF *Rhizophagusirregularis*, which is called the *MAT* locus in other sexual fungi. I employed bioinformatic, population genetic, comparative genomic, and experimental approaches to answer the specific questions: (1) do AMF harbour genes which are normally found within the *MAT* locus of other fungi, (2) does allelic variation exist at these genes which resembles variation found at such genes orthologues in other fungi, (3) does syntenic conservation exist at one of these genes, and (4) during crosses of AMF isolates, do expression patterns of these genes resemble such patterns found at gene orthologues during crosses of other fungi.

Chapter 1:

Searching for clues of sexual reproduction in the genomes of arbuscular mycorrhizal fungi

Rohan Riley and Nicolas Corradi

RR performed the literature review and wrote the initial draft of the manuscript while NC provided critical oversight of the writing and editing process.

Canadian Institute for Advanced Research, Department of Biology, University of Ottawa, Ottawa, ON, Canada, K1N 6N5

Abstract

Arbuscular mycorrhizal fungi (AMF) represent an ecologically and evolutionarily important group of land plant symbionts, which produce multi nucleated spores and hyphae that are currently thought to have propagated clonally for over 500 million years. This long-term absence of sex in AMF is a puzzling evolutionary feature that has sparked scientific interest for some time, but a provoking explanation for their successful evolutionary history in the absence of an obvious sexual cycle is that these organisms may have cryptic sex, or a parasexual life cycle, allowing them to recombine alleles and compensate for deleterious mutations. Interestingly, the recent acquisition of large sequence data from many AMF species can finally allow this hypothesis to be tested more extensively. In this perspective, we highlight emerging evidence based on sequence data for the potential of AMF to have sexual reproduction, and propose a number of routes that could be taken to further explore the presence (or absence thereof) of sex in this poorly studied, yet highly relevant, fungal group.

Introduction

Arbuscular Mycorrhiza Fungi (AMF) are an ancient and ubiquitous group of fungi that form an ecologically important symbiosis with over 80% of land plants and many economically important crops; also known as the mycorrhizal symbiosis (Bonfante and Genre 2010; Corradi and Charest 2011; Sanders 2002; Sanders 2003). A hallmark of this symbiosis is the substantial expansion of the plant rhizosphere by the fungal mycelium, which increases the transfer of nutrients from the surrounding soil to the plant, and results in a significant contribution to the global productivity of terrestrial ecosystems. AMF are also known to harbour a number of cellular and genetic features that are at odds with those found in other eukaryotes. These include the presence of coenocytic cells that are perpetually multinucleated, resulting in hundreds of nuclei being present simultaneously in one cytoplasm. Genetic polymorphism in AMF has sparked extensive debate about the genetic organization of these nuclei (heterokaryosis *vs* homokaryosis) (Hijri and Sanders 2005; Kuhn *et al.* 2001; Pawlowska 2005; Pawlowska and Taylor 2004; Sanders and Croll 2010), but the consensus view is that certain genes can sometimes exhibit significant polymorphism within one individual (Hijri *et al.* 1999; Hijri and Sanders 2005; Kuhn *et al.* 2001; Lanfranco *et al.* 1999; LloydMacgilp *et al.* 1996; Pawlowska and Taylor 2004; Rodriguez *et al.* 2004).

AMF are also very intriguing from an evolutionary point of view, as these fungi are currently thought to have propagated for over 500 million years (Humphreys *et al.* 2010; Redecker *et al.* 2000) in the absence of sexual reproduction - a very rare feature that has placed AMF in a “highly select” group of organisms called “ancient asexuals”, along with a few, distantly related lineages (*e.g.* bdelloid rotifers, darwinulids) (Butlin *et al.* 1998; Smith 1986; Welch and Meselson 2000). The long-term absence of sex in these organisms is, indeed, highly unusual in eukaryotes, because its presence is in sharp conflict with the evolutionary view that such lineages should accumulate many deleterious mutations and become quickly extinct (Judson and Normark 1996; Normark *et al.* 2003). In the case of AMF, explanations for this exclusively clonal reproduction have usually centered on their atypical multinucleated state, and especially on the potential heterokaryotic nature of their cells (*e.g.* simultaneous presence of genetically diverse nuclei in one cytoplasm) (Hijri and Sanders 2005; Kuhn *et al.* 2001). This hypothesis, however, has been challenged (Pawlowska 2005; Pawlowska and Taylor 2004), so the atypical longevity of AMF may also be explained by other modes of propagation that are currently unknown.

In particular, an alternative explanation for the evolutionary success of AMF is that they may be able to undergo cryptic sex; or a parasexual life cycle. Indeed, many cases exist where asexual lineages have been ultimately found to undergo a sexual cycle; including a number of important pathogens such as *Giardia* spp. (Ramesh *et al.* 2005) and members of different fungal lineages (e.g. *Candida albicans*, *Aspergillus fumigatus*) (Dyer and O'Gorman 2011; Hull and Heitman 2002; Hull and Johnson 1999; Hull *et al.* 2000; O'Gorman *et al.* 2009; Wong *et al.* 2003), some of which exhibit atypical sexualities such as the same-sex mating of *Cryptococcus neoformans* and *Candida albicans* (Alby *et al.* 2009; Lin *et al.* 2005). So are AMF really asexual? In this perspective we review our current knowledge regarding the presence in AMF genomes, of events and genes that are typically linked with the presence of sex in other eukaryotes (Dunthorn and Katz 2010; Malik *et al.* 2008; Parfrey *et al.* 2008; Schurko and Logsdon 2008; Schurko *et al.* 2009). Finally, we propose a number of new, speculative routes that could be used to further explore the presence (or absence) of sexuality (or parasexuality) in this poorly studied fungal lineage.

Known signatures of sex in AMF genomes

Recombination

An essential mechanism involved in sexual reproduction is meiotic recombination; a process that plays a role in shuffling homologous alleles and in reducing the amount of deleterious mutations carried within one nucleus. Its presence in an organism can be investigated by comparing sequences of homologous loci between different individuals, or by identifying the genes involved in the “core" meiotic machinery. Both approaches have been recently targeted to sequence data from AMF, revealing that similar events have happened in the genomes of these putatively ancient asexual species.

To date, presence of recombination events in AMF populations have been investigated using a variety of statistical tests, all targeted to multi-locus sequence typing data from members of natural populations and different species (*Glomus claroideum, Glomus irregulare, Glomus etunicatum*) (Croll and Sanders 2009; den Bakker *et al.* 2010; Vandenkoornhuyse *et al.* 2001). Surprisingly, all these studies were able to detect footprints of recombination using a number of bioinformatics tools and models, suggesting that these events are either presently occurring in natural populations of AMF, or that they did occur at some point. Since meiosis has never been observed in AMF, however, it has been impossible to determine whether these recombination events have resulted from sexually-related processes (*e.g.* meiosis), or from independent asexual mechanisms (*e.g.* mitosis, transposition). A recent study has nevertheless demonstrated that AMF could undergo a conventional meiosis; if they really wanted to (Halary *et al.* 2011).

Potential for meiosis in AMF

Meiosis is a tightly controlled activity that requires the expression of a conserved “core” of genes in order to fully function. These genes are not supposed to be present and maintained in the genomes of “ancient asexuals”, because the loss of their original functions would be expected to result in their “pseudogenization” over such long periods of time (unless these gene would evolve an unlikely, completely new function). Recent explorations of large amounts of AMF genome and transcriptome data have, however, found that the genomes of many AMF species harbour a conserved and expanded set of meiosis-related genes; whose complement is typical of that of many other sexual fungi (Halary *et al.* 2011). In particular, four species in the genus *Glomus* have all been found to harbour proteins involved in the “core meiotic machinery”, including 7 gene orthologues that are known to function exclusively in meiosis in many other eukaryotes (*Rec8*, *Spo11*, *Mnd1*, *Hop2*, *Dmc1*, *Msh4*, *Msh5;* Table 1). The identification of this gene set in the AMF genome is a highly relevant finding that suggests meiosis is likely to occur (in one form or another) in these supposedly ancient asexuals, and that it may account for the presence of recombination events previously reported in different AMF species. A small portion of these latter genes were found to be transcribed at different stages of the life cycle of *Glomus intraradices* (Tisserant *et al.*), suggesting that some may be expressed continually in these symbiotic fungi.

AMF harbor homologues of sexP and sexM from the zygomycetes

The presence of recombination across many AMF genomesand their possible meiotic origin, suggest that these ecologically critical fungi may be capable to either undergo some sort of “hidden” sexuality, or a parasexual life cycle. If this is the case, however, one

Table 1. List of "Core meiotic" genes from *Saccharomyces cerevisiae* found to function exclusively in meiosis based on (Halary *et al.* 2011). The protein name, accession number, and function of the proteins are listed.

Protein	***Accession#***	***Description***
Rec8	**CAY86963**	Meiosis-specific component of sister chromatid cohesion complex; maintains cohesion between sister chromatids during meiosis I; maintains cohesion between centromeres of sister chromatids until meiosis
Spo11	**YHL022C**	Meiosis-specific protein that initiates meiotic recombination by catalyzing the formation of double-strand breaks in DNA via a transesterification reaction; required for homologous chromosome pairing and synaptonemal complex formation
Mnd1	**YGL183C**	Protein required for recombination and meiotic nuclear division; forms a complex with Hop2p, which is involved in chromosome pairing and repair of meiotic double-strand breaks
Hop2	**YGL033W**	Meiosis-specific protein that localizes to chromosomes, preventing synapsis between nonhomologous chromosomes and ensuring synapsis between homologs; complexes with Mnd1p to promote homolog pairing and meiotic double-strand break repair
Dmc1	**YER179W**	Meiosis-specific protein required for repair of double-strand breaks and pairing between homologous chromosomes; homolog of Rad51p and the bacterial RecA protein
Msh4	**YFL003C**	Protein involved in meiotic recombination, required for normal levels of crossing over, colocalizes with Zip2p to discrete foci on meiotic chromosomes, has homology to bacterial MutS protein
Msh5	**YDL154W**	Protein of the MutS family, forms a dimer with Msh4p that facilitates crossovers between homologs during meiosis; msh5-Y823H mutation confers tolerance to DNA alkylating agents; homologs present in C. elegans and humans

should expect other sexually important genes to be present in their genomes; including most of those that are typically involved in sexual recognition and mating in other fungi (Heitman 2006; Heitman 2010; Poggeler 2002). Fungal sexuality has been demonstrated to be governed by a particular genomic region commonly referred to as the fungal mating type (*MAT*) locus (Heitman 2010; Lee *et al.* 2010). The genes located at this particular locus are usually conserved in their order among members of the same fungal group, so it should be relatively straightforward to search for their presence and order across the sequence data that is currently available from AMF. This approach has been recently applied to search for potential *MAT*-related genes across the transcriptome of *Glomus intraradices* (Tisserant *et al.*).

Importantly, these searches resulted in the identification of some genes that strongly resemble those found within the *MAT* loci of many zygomycetes, a basal fungal and paraphyletic phylum from which the AMF may have originated (Corradi and Sanders 2006; Halary *et al.* 2011; Lee *et al.* 2009; Liu *et al.* 2009; Liu *et al.* 2006). These genes encode for orthologs of the sexM and sexP proteins from the zygomycetes (e.g. *Phycomyces* spp., *Mucor* spp., *Rhizopus* spp., and *Syzygites* spp.) (Gryganskyi *et al.* 2010; Idnurm 2011; Idnurm *et al.* 2008;Lee *et al.* 2008); two highly divergent alleles that encode for high mobility group proteins (HMG) associated with sexual identity in this fungal lineage. In the zygomycetes, these alleles have been found surrounded by genes encoding an RNA helicase on one side, and either a triose phosphate transporter or a glutathione oxidoreductase on the other (Idnurm 2011). Unfortunately, the sequence data available from AMF (e.g. transcriptome data or short genomic contig) is currently too restricted and short to identify which genes are flanking the AMF homologues of *sexP* and *sexM*, but the identification of these genes in AMF represents a very exciting finding that certainly warrants further investigation in this specific area of AMF genomic research. Certainly, efforts should now focus specifically on identifying which genes surround the homologues of *sexM* and *sexP* in the genomes of AMF, as their discovery may reveal the presence of gene order conservation with the *MAT* loci of other fungi.

Further exploring cryptic sex in AMF

Current evidence for sexuality in AMF is highly intriguing but still poor

In the last few years, a number of studies have gathered evidence for the presence of recombination and meiosis in AMF, as well as the presence of AMF genes that strongly resemble those composing some fungal *MAT* loci. The acquisition of this information represents obvious breakthroughs in our understanding of the evolution of AMF, as it suggests that these organisms may not be the evolutionary aberration that they have long held to be. However, current knowledge about sex in AMF is still lacking at many levels, and a number of studies still need to be carried out before we will be able to tell with certainty whether AMF can, or cannot, undergo sexual, or parasexual, life cycles.

For instance, all studies of recombination in AMF have pointed towards a rare occurrence of these events, but it is currently unknown whether the low frequency of such events reflects a rarity of recombination across AMF genomes, or a side effect of poor sampling and sequencing strategies (i.e. little sequence data from relatively few individuals have been analyzed so far). In order to clear up this potential controversy, future investigations in the field should search for recombination across a larger sequence data and, ideally, compare such data among many more AMF individuals. Accordingly, the presence of AMF orthologues of meiosis-specific genes cannot be taken yet as conclusive evidence that meiosis is occurring in these fungi, because this mechanism has never been observed in AMF and it is virtually possible (though very unlikely) that these organisms may have evolved their own way to use this machinery without having to undergo a typical meiosis.

Certainly, experiments using microscopy technologies and heterologous expression vectors will be required to determine if AMF genes can, at least, restore a conventional meiosis in model organisms (*e.g.* using *S. cerevisiae* with meiosis-specific gene knocked-out). AMF have also been observed to undergo hyphal fusion (anastamosis), and this has been linked with the potential presence of nuclear exchange in several studies (Angelard and Sanders 2011; Croll *et al.* 2009; de la Providencia *et al.* 2005; Giovannetti *et al.* 1999; Koch *et al.* 2006). This cytoplasmic bridging is used by many other fungi to exchange genetic material following sexual recognition, so studying the biochemical causes and consequences of anastomoses in AMF could be a key for understanding how these organisms recognize each other and, perhaps, undergo nuclear exchange (i.e. the very first step of sexual reproduction in many fungi).

Perhaps, the hardest task in the search for AMF sexuality will be the identification of a potential *MAT* locus in AMF, because genome information from this group is still scarce, and the *MAT* loci are known to vary greatly in shape and form across members of different fungal phyla. The acquisition of a large amount of genome sequence data from these organisms appears, therefore, essential for this type of exploration, and likely the only possible route that could result in the identification of synteny conservation between "*MAT*-gene candidates" from AMF and those of other fungal lineages. For example, the genome sequencing project of *Tuber melanosporum* revealed the presence of a *MAT* loci in this ectomycorrhizal fungus, thus changing our current views on the sexuality of this economically important organism (Martin *et al.* 2010). Finally, presence of substantial allelic variation should also be searched for across potential *MAT* gene candidates from different AMF individuals (Table 2) because similar variation has been found to govern sexual identity in many other fungi.

Which pathways, directly or indirectly involved in sexuality, are left to discover in AMF?

In other fungi, sexual reproduction involves the interaction of complex pathways to initiate and terminate mating processes between different fungal individuals. Some of these pathways are involved in the pheromone dialogue between sexually distinct members of one species, and are often required to trigger partner recognition, hyphal fusion (anastomosis) and subsequent nuclear exchange between partners. Unfortunately, the biochemical processes that control the dialogue between potential fungal mating partners are very diverse between different fungal groups.

For instance, the genes producing intermediates of trisporic acid are central for the development of sexual hyphae in a number of zygomycetes (Burmester *et al.* 2007), and because most recent phylogenetic analyses pointed towards a potential similarity of AMF with members of this group (Corradi and Sanders 2006; Halary *et al.* 2011; Lee *et al.* 2009; Liu *et al.* 2009; Liu *et al.* 2006), AMF could also use orthologues of such genes for similar purposes. In the trisporic acid synthesis pathway, intermediates produced in each mating type are delivered to opposite mating types to trigger the production of trisporic acids, the pheromone triggering the formation of specialized hyphae leading to sexual spore production (Lee *et al.* 2010). Other genes involved in sexual cell formation include those composing the MAPK signal transduction cascade, a pathway that has no known link with the zygomycete trisporic acid pathway, but is essential for mating partner recognition in the model ascomycete, *S. cerevisiae* and in the basidiomycete, *Cryptococcus neoformans* (Jones and Bennett 2011; Rispail and Di Pietro 2010) (See table 2 for a list of MAPK and trisporic acid pathway components).

Genes involved in these latter pathways should also be searched for across the genomes of different AMF, as these may represent a stepping stone towards understanding how AMF individuals recognize each other and perhaps, subsequently exchange genetic material (Croll *et al.* 2009). Ultimately, tracking the fate of nuclei following anastomosis will result in a major breakthrough in our understanding of the origin and evolution of this critical fungal group. Future studies in this area are especially warranted, as they could reveal whether some nuclei are prone to undergo nuclear fusion following exchange between genetically different AMF individuals, resulting perhaps in diploid nuclei that could readily undergo meiosis.

Table 2. List of proteins known to be involved in a number of pathways that are directly or indirectly related with the presence of sexual reproduction in many fungi. Proteins involved in fungal mating, including the yeast pheromone response MAPK cascade, the Zygomycete trisporic acid pathway, and genes associated with the *MAT* locus of the Zygomycetes. Accession numbers and respective functions are listed. Species codes are *Blakeslea trispora* (Bt), *Mucor mucedo* (Mm), *Phycomyces blakesleeanus* (Pb), *Saccharomyces cerevisiae* (Sc), and *Syzygites megalocarpus* (Sm). N.A.; not available

Protein	**Accession #**	**Species**	**Function**
MAPK Cascade			
Ste3	**NP_012743**	Sc	Mating-type "a" G-protein coupled receptor which binds α-factor pheromone and acts as a guanine exchange factor for Gpa1 .
Ste2	**NP_116627**	Sc	Mating-type "α" equivalent of Ste3. A G-protein coupled receptor which binds a-factor pheromone.
Gpa1	**NP_011868**	Sc	G-protein α-subunit. Becomes activated by Ste3 or Ste4, depending on mating type, and subsequently decouples from β and γG-protein subunits
Ste4	**NP_014855**	Sc	G- protein β-subunit. Decouples from α-subunit as a heterodimer with γ-subunit and interact with Ste20, Ste5, and Far1.
Ste18	**NP_012619**	Sc	G-protein γ-subunit. Decouples from α-subunit as a heterodimer with β -subunit .
Cdc42	**NP_013330**	Sc	Rho-like GTPase; Activated by Cdc24 and activates Ste20.
Cdc24	**NP_009359**	Sc	Likely activated by Far1 and serves as a guanine nucleotide exchange factor for Cdc42.
Far1	**NP_012378**	Sc	Activated by the MAPK Fus3 and subsequently inhibits the cell cycle.
Ste20	**NP_011856**	Sc	MAPKKKK which becomes activated by Cdc42-GTP and subsequently activates the MAPKKK, Ste11.
Bem1	**NP_009759**	Sc	Membrane associated protein which binds to Ste20, bringing Ste20 in high local concentration with other cascade components.

Ste5	**NP_010388**	Sc	Scaffold protein which associates with the cell membrane and is essential for the pheromone response in yeast; binds all three kinases (Ste11, Ste7 and Fus3).
Ste11	**NP_013466**	Sc	MAPKKK, activated by Ste20.
Ste7	**NP_010122**	Sc	MAPKK, activated by Ste11
Fus3	**NP_009537**	Sc	MAPK, activated by Ste7.
Ste12	**NP_011952**	Sc	Transcription factor activated by the MAPK, Fus3. Co-regulated by binding proteins to elicit mating and invasive growth responses.
Dig1	**NP_015276**	Sc	Together with Dig2 the complex interacts with Ste12, blocking its activity. Complex dissociates following phosphorylation of Ste12.
Dig2	**NP_010768**	Sc	Forms a complex with Dig1 to regulate Ste12.
Mcm1	**NP_013757**	Sc	Interacts with Ste12 and Matα1 to specify mating target gene expression.
Fus1	**NP_009903**	Sc	Proposed coordinator of signaling, fusion and polarization events required for cell-cell fusion.
Production of trisporic acids			
Tsp3	**CAL64769**	Bt	Carotene oxygenase which performs the initial cleavage of β-Carotene necessary for the production of intermediates involved in the chemical "cross-talk".
Tsp2	**AM937248.1**	Mm	4-dihydrotrisporin dehydrogenase; converts 4-dihydrotrisporin to trisporin in the (-) mating type which is delivered to the (+) mating type
Tsp1	**Q01213**	Mm	4-dihydromethytrisporate dehydrogenase; converts 4-dihydromethyltrisporate, received from the (+) mating type, to methytrisporate in the (-) mating type
***MAT* locus of the Mucorales**			
SexM	**ABX27909.1**	Pb	High Mobility Group transcription factor characterizing the (-) mating type MAT locus of *Phycomyces blakesleeanus.*

SexP	**ABX27912.1**	Pb	High Mobility Group transcription factor characterizing the (+) mating type MAT locus of *Phycomyces blakesleeanus.*
TPT	**ABX27911**	Pb	Triose phosphate transporter. Located in the MAT locus of *Phycomyces blakesleeanu,* flanking SexM or SexP.
RNA Helicase	**ABX27913**	Pb	RNA helicase located in the Mat locus of *Phycomyces blakesleeanus,* flanking sexM or SexP.
Glutathione reductase	**N.A**	Sm	Glutathione oxidoreductase located in the MAT locus of *Syzygites megalocarpus* flanking sexM of sexP

Chapter 2:

Identification and analysis of an expanded family of fungal *MAT*-HMG genes in the arbuscular mycorrhizal fungus *Rhizophagus irregularis*

Rohan Riley[1], Philippe Charron[1], Alexander Idnurm[2], Francis Martin[3], Yolande Dalpé[4] and Nicolas Corradi[1]

RR performed all QPCR procedures and data analysis while PC assisted with conventional PCR procedures. AI performed inverse PCR procedures, FM provided the transcriptome and genomic datasets of *Rizophagus irregularis* isolate DAOM 197198, and *Rhizophagus intraradices* isolate DAOM 181602. YD provided genomic DNA of all DAOM isolates used for the population screening while NC was the principal investigator of the study, providing important oversight in the data acquisition and analysis process.

[1]Canadian Institute for Advanced Research, Department of Biology, University of Ottawa, Ottawa, ON, Canada, K1N 6N5

[2]Division of Cell Biology and Biophysics, School of Biological Sciences, University of Missouri-Kansas City, Kansas City, Missouri, United States of America

[3]UMR INRA-UHP 'Interaction Arbres/Micro-Organismes', Centre INRA de Nancy, Champenoux, France

[4]Agriculture and Agri-Food Canada, 960 Carling Ave. Ottawa, On, K1A 0C6, Canada

Abstract

Arbuscular mycorrhizal fungi (AMF) are ecologically important symbionts of land plants, which have been considered evolutionary anomalies due to a long-term lack of an observable sexual cycle. However, recent investigations across available sequence data from AMF have revealed the presence of homologues of fungal *MAT*-High mobility group (HMG) and core meiotic genes. This challenges the overall notion that AMF have exclusively evolved clonally for over 500 million years. By exploring a large-scale transcriptome data set from three AMF species, we demonstrate that *MAT*-HMG genes are present in inflated numbers in

AMF, and we describe the evolution of this gene family across different isolates of *Rhizophagus irregularis*. Interestingly, we found that some HMGs harboured allelic polymorphism between isolates whose nature mirror that typically found between known fungal mating type loci. Exploration of this sequence dataset also allowed for the identification of the mutational events that affected the evolution of this gene family, which include gene conversions detected between gene paralogues in a single isolate and gene orthologues contained between isolates. Selected members of this gene family were used for downstream functional analyses of gene expression, using Real-Time PCR on cDNA generated from different crossing-experiments with genetically variable isolates of *R. irregularis*. These procedures suggested some *MAT*-HMGs may be involved in partner recognition in these relevant fungi. Finally, a unique genomic structure composed of at least three tandem-repeated *MAT*-HMG genes is also reported.

Introduction

Arbuscular Mycorrhizal Fungi (AMF) are an ancient and ubiquitous group of obligate plant symbionts that are thought to have assisted the colonization of land by plants approximately 500 million years ago through the establishment of the mycorrhizal symbiosis. This symbiosis involves the close association of AMF with the roots of over 80% of land plants, and many important crops, including *Zea mays* (Corn), *Glycine max* (Soya bean), *Triticum aestivum* (Wheat), *Solanum tuberosum* (Potato)and many others (Bonfante and Selosse 2010; Corradi and Charest 2011; Sanders 2002; Sanders 2003; Smith *et al.* 1997).The hallmark of this symbiosis is the transfer of photosynthetically produced carbohydrates from the plant to the fungus in exchange for increased supplies of water and mineralized nutrients;

most notably phosphates and nitrates. In particular, the AMF benefit the plants by extending their hyphae towards soil areas that are beyond the reach of the roots, allowing the plant to access nutrient zones that they could not obtain otherwise (Smith *et al.* 1997).The presence of AMF diversity in soil has also been associated with increased above ground plant and below ground microbial diversity (Munkvold *et al.* 2004; Smith *et al.* 2000; van der Heijden *et al.* 1998; Wagg *et al.* 2011), and for these reasons AMF are considered keystone microbial organisms in terrestrial ecosystems;contributing significantly to global productivity and biodiversity.

AMF are intriguing from a cellular point of view, as their hyphae are ceonocytic (i.e. no cellular separation) and perpetually multinucleate which allow nuclei and other cytoplasmic components to move freely throughout one common cytoplasm (Rosendahl 2008). These multinucleated hyphae can sometimes fuse through a process called anastomosis, at times resulting in cytoplasmic contact between genetically different individuals (Croll *et al.* 2009). Genetic diversity appears to be an hallmark of these fungi, as anumber of studies have identified the presence of extensive molecular polymorphism within single AMF spores. However, the origin of this variation is currently unknown, as the distribution of this genetic variation has been proposed to either result from extensive genome duplications (homokaryosis and polyploidy), or tobe a consequence of genetically different nuclei co-existing within one cytoplasm (heterokaryosis) (Hijri and Sanders 2005; Kuhn *et al.* 2001; Lanfranco *et al.* 1999; LloydMacgilp *et al.* 1996; Pawlowska and Taylor 2004; Rodriguez *et al.* 2004).

From an evolutionary point of view, AMF have been considered to be somewhat of a scandal as they are currently thought to havepropagated for the past 500 million years in the absence of sexual reproduction (Judson and Normark 1996). Specifically, exclusively asexual lineages are predicted to accumulate deleterious mutations and rapidlygo extinct (Judson and Normark 1996; Normark *et al.* 2003), so the extreme longevity of AMF resulted in them being referred to as "ancient asexuals"; a label that they share with a handful of distantly related lineages such as the bdelloid rotifers and darwinulids (Butlin *et al.* 1998; Smith 1986; Welch and Meselson 2000).

Importantly, many organisms that were long thought to have evolved only clonally have been ultimately found to either undergo sexual reproduction (i.e. *Giardia,(Logsdon Jr 2008)*), or to harbour several genes that are only known to function in sexually-related processes. These taxa include several fungal lineages of medical and industrial importance (i.e. *Aspergillus fumigatus*, *Aspergillus oryzae,Candida albicans* (Bennett 2010; O'Gorman *et al.* 2009;Wada *et al.* 2012)), and it is therefore intriguing to speculate about the potential presence of cryptic sex in the AMF. In fact, genome explorations across available sequence data from AMF have also revealed some genetic signatures that are generally linked with sexual reproduction, including the presence of events of recombination within several natural populations of AMF (Croll and Sanders 2009; den Bakker *et al.* 2010; Vandenkoornhuyse *et al.* 2001), and the identification of severalAMF genes whose homologues are otherwise associated with fungal sexuality (Riley and Corradi 2013). These genes include the core meiotic tool-kit (Halary *et al.* 2011), a suite of genes that is essential to initiate and complete meiosis, and the presence of homologues of *SexM* and *SexP*, which are proteins normally found in the Mating-type (*MAT)* loci ofearly diverging fungi in the phylum Zygomycota (Tisserant *et al.* 2012).

The fungal *MAT* locus is a genomic region found in most fungi which serves the common purpose of determining sexual compatibility between two individuals of a species. Sexual compatibility between individuals is determined by allelic variation at this locus, and the gene content and structure of the locus can differ quite drastically across members of the fungal Kingdom (Lee *et al.* 2010).Alleles of the *MAT* locus in any species are often called idiomorphs to denote sequences which occupy the same locus on a chromosome, but do not necessarily share conservation in gene order or a common descent (Metzenberg and Glass 1990). Overall, the *MAT* loci of different fungi can contain different combinations of genes, which typically include homeodomain, HMG transcription factors, or alpha-box which has recently been reclassified as an HMG domain (Martin *et al.* 2010).The oldestversion of the fungal MAT locus is known from members of the zygomycetes, where the two idiomorphs are known as *SexM* or *SexP* (Idnurm *et al.* 2008). Along these idiomorphs, substantial allelic variation occurs withinone gene encoding for a high mobility group protein (HMG) of the MATA_HMG subfamily, (Martin *et al.* 2010).This latter subfamily is only found in the fungi, and is currently thought to represent the ancestral transcription factor of fungal *MAT* loci (Idnurm 2011; Idnurm *et al.* 2008;Lee *et al.* 2008).

In the zygomycetes, the function and target genes of the *SexM* and *SexP* transcription factors are not known, however several lines of evidence implicate them in partner recognition and sexual development. For instance, in the zygomycete *Mucor cicinelloides,* S*exM* mutants shows no sexual development when genetically different individuals are crossed (i.e. crossed cultures) (Li *et al.* 2011),and in *Phycomyces blakesleeanus* crossings involving individuals of different mating-types, the *SexM* and *SexP* genes are upregulated (Idnurm *et al.* 2008; Wetzel *et al.* 2012). Furthermore, *P.blakesleeanus* individuals harbouring both

SexM and *SexP* were found to have the capacity to be homothallic (i.e. capable of self-reproduction) (Idnurm *et al.* 2008). Altogether, these findings support the important role of *SexM* and *SexP* genes in sexual differentiation and partner recognition in the Zygomycota.

In AMF, homologues of *SexM* and *SexP* have been identified, but their overall evolution within AMF genomes and their potential role in partner recognition or sexual reproduction are currently unknown. In this study, we surveyed the transcriptomes of two isolates of *Rhizophagus irregularis*(DAOM197198 and isolate C2 (referred to as SwiC2 in this study), the latter isolated from a field in Tänikon, Switzerland and named according to (Koch *et al.* 2006), *Rhizophagus intraradices* (DAOM 181602) and one isolate of *Rhizophagus diaphanus* (MUCL 43196) for the presence of homologues of *SexM* and *SexP*. These explorations revealed the presence of a surprisingly elevated number of gene homologues in all strains.In parallel, the potential presence of idiomorphs of these AMF *MAT*-HMGs was assessed by isolating orthologues of all genes from otherstrains of *R. irregularis* originating from various populations, followed by an assessment of the differential expression of a subset of these genes using Real-Time PCR proceduresfollowing isolate interaction experiments. Finally, investigations of genomic regions surrounding some of these genes revealed an intriguingly unique genomic structure, which may reflect the structure of other regions of the AMF genome.

Materials and Methods

AMF culture and tissue isolation for DNA extraction

In the present study, a total of 20 isolates (Table 1) from the AMF *Rhizophagus irregularis*, one isolate of *R. intraradices* (DAOM 181602), and *R. diaphanus* (MUCL 43196) have been

Table 1.Strains of*Rhizophagusirregularis* used in this study.

DAOM	Origine
240277	Poland (Pol)
197198	Pont-Rouge Québec Canada (CanQc1)
234180	Ripon Québec (CanQc2)
234181	Îles-de-la-Madeleine Québec (CanQc3)
240201	Îles-de-la-Madeleine Québec (CanQc4)
240448	Tunisia (Tun)
240434	Larose Forest Ontario (CanOn1)
240721	Belgium, Louvain-la-Neuve (Bel)
229457	Clarence Creek Ontario (CanOn2)
240159	Revelstoke BC (CanBc)
234328	Scandinavia (Sca)
A1	Switzerland (SwiA1)
A2	Switzerland (SwiA2)
A4	Switzerland (SwiA4)
A5	Switzerland (SwiA5)
B3	Switzerland (SwiB3)
C2	Switzerland (SwiC2)
G	Switzerland (SwiG)
H	Switzerland (SwiH)
141	Switzerland (Swi141)

investigated at different levels for the presence of potential *MAT*-HMG genes. Several isolates of *R. irregularis* (referred here as SwiA4, SwiC2, SwiB3, SwiG, SwiH, SwiA1, SwiA5, SwiA2, Swi141) were previously harvested from one field in Switzerland, and cultivated within *in-vitro* split-plates in symbiotic association with Ri T-DNA-transformed *Daucus corrota* roots as previously described (Koch *et al.* 2006). The isolates SwiA4, SwiC2 and SwiB3 were chosen for preliminary assessment of genetic variation in this study because they are relatively easy to propagate under axenic conditions, and have been genetically investigated by others in a number of instances (Angelard and Sanders 2011; Corradi *et al.* 2007; Croll *et al.* 2009; Croll *et al.* 2008; Koch *et al.* 2006), have been repeatedly found to be genetically dissimilar despite being from the same population (Corradi *et al.* 2007; Croll *et al.* 2008), and have also been found to be capable to undergo anastomosis (Croll *et al.* 2009).

AMF are grown in symbiosis with Ri T-DNA-transformed *Daucus corrota* roots using petri-dishes separated in two compartments (referred here as split-plates). The use of split plates allows the separation of AMF mycelium from the host, ensuring that only root-free AMF mycelia and spores are harvested for downstream procedures. M-medium containing root-free hyphae and spores were extracted from the split plates and dissolved in 10mM citrate acid/ citrate buffer solution within 3L beakers and stirredwith a magnetic stir rod for 1.5 hours. Following M-media dissolution, AMF hyphae were harvested using a sieve and washed twice with ddH_2O. Water was then vacuumed out of the fungal mycelia using an in-house system and either processed immediately or frozen at $-80^{o}C$.

Genomic DNA extraction

Genomic DNA was extracted using variable amounts of mycelium usingour “in house” protocol based on phenol-chloroform extraction with the MasterPure^{TM} Complete DNA and RNA purification kit from Epicentre Biotechnologies (Madison, WI). Briefly, mycelia was placed in 300µL of lysis solution containing 1µL of proteinase K, crushed using a plastic pestle and immediately incubated at 65^{o}C for 15 minutes, vortexed every 5 minutes. Samples were then cooled to 37^{o}C and 2µL of 5µg µL^{-1}RNase A was added to the mixture, which was then incubated 30 minute at 37^{o}C. 600µL of phenol was placed in the solution which was then vortexed for 5 minutes and centrifuged at 4^{o}C at 10 000 ***g***for 10 minutes to pellet debris. The supernatant was placed into a new tube, and followed by the addition of 600µL of chloroform, vortexing for 5 seconds, and then centrifugation at 4^{o}C at 10 000 ***g*** for 10 minutes to pellet debris. The supernatant was placed in a new tube and DNA was precipitated using 99% isopropanol following an incubation for 30 minutes at -20^{o}C, and a final centrifugation at 4^{o}C at 10 000 ***g*** for 10 minutes. The supernatant was discarded and the DNA pellet was washed twice with 70% v/v Ethanol, vacuum dried and resuspended in TE buffer.

RNA extraction

Total RNA was isolated using the QIAgen Plant RNA extraction kit. Between 20 and 40 mg of fresh mycelium was crushed in β-Mercaptoethanol diluted with the manufacturers lysis buffer using a plastic mortar.The homogenized mycelium was then subjected to RNA extraction following manufacturer's instructions. Upon RNA extraction, the solution was treated with RNase-Free DNase I (Epicentre Biotechnologies) at 37^{o}C for one hour followed by DNase I treatment in the presence of 200 µL of T and C Lysis solution (Epicentre Biotechnologies). The resulting solution was vortexed for 5 seconds, 200 µL of MPC protein precipitation reagent (Epicentre Biotechnologies) were added for protein precipitation, and the mixture was vortexed for an additional 5 seconds. The solution was placed on ice for 5 minutes, followed by a 10 000 ***g***centrifugation at 4^{o}C for 10 minutes to pellet debris. Supernatant was placed into a new tube, and RNA was precipitated using isopropanol, washed with 70% v/v ethanol,vacuum dried, and ultimately re-suspended in RNase-free H_2O (Epicentre Biotechnologies). RNA concentration was determined using a Nanodrop spectrophotometer (Fisher Scientific). In all cases, 1µg of DNase-free RNA was immediately subjected to RT-PCR using iScript kit (Bio-Rad Laboratories) following the manufacturers protocol. In addition, 0.5µg of RNA was heated at 65^{o}C for five minutes then ran on a 2% agarose 1X Tris-EDTA gel to visually inspect RNA forconsistent quality across samples. No-RT controls were performed alongside all cDNA syntheses reported in the present study to confirm the absence of genomic DNA contamination.

Acquisition of transcriptome data from R. irregularis and R. diaphanousand identification of AMF MATA_HMG domains

RNA isolated from *in-vitro* cultures of *R. irregularis* and *R. diaphanous* was subjected to Illumina sequencing at a facility located in Geneva, Switzerland (Fasteris S.A.). The respective transcriptomes were used to produce cDNA libraries following the company’s protocols, which were then sequenced using one complete channel on the HiSeq2000 instrument, using and 100 bp-paired reads. The sequencing procedure resulted in 201'051'108 and 176'382'504 reads for *R. diaphanus* and *R. irregularis*, respectively. Reads were assembled using Velvet Oases and a hash value of 93 for both species, which resulted in the acquisition of approximately 20,000 contigs for each species.

Homologues of fungal mating type HMG proteins representative of different fungal phyla (*MAT*-HMG; n=25) where searched across available transcriptome sequence data newly obtained from *R. irregularis* (strain SwiC2; this study, *R. diaphanus* (MUCL 43196; this study) and *R. irregularis* (DAOM 197198, (Tisserant *et al.* 2012)) using reciprocal Blast procedures (i.e. BlastX, tBlastX, BlastP, tBlastN). Similar searches were also performed across publicly available genome sequence data from members of the Chytridiomycota, Zygomycota, Basidiomycota, Ascomycota (i.e. the chytrids *Allomyces macrogynus,* and B*atrachochy triumdendrobatidis ;* the basidiomycete*s Ustilago maydis, Puccinia graminis* and *Cryptococcus neoformans;* the ascomycetes *Saccharomyces cerevisiae, Aspergillus nidulans and Neurospora crassa;* and the zygomycetes *Phycomycesblakesleeanus*and *Rhizopus oryzae*). All potential *MAT*-HMGs identified in the transcriptomes of *Rhizophagus* and other fungal genomes were further compared against the nr database in order to confirm their homology, and their sequences were manually inspected to avoid redundancy.

qPCR procedures

Primers for qPCR were designed to amplify a small region of 6 selected HMG sequences, oneexternal control(β-Tubulin), and an internal reference gene(Ef1α) (Table 2). Primers were designed using primer3 (Rozen and Skaletsky 2000), with the exception of the locus HMG6 where two primer sets were designed manually over polymorphic regions corresponding to isolate-specific alleles of the locus; HMG6 (i.e. the allele HMG6-A being specific to isolate SwiA4, and the allele HMG6-B being specific to isolate SwiC2).

Real time PCR reaction were carried out on a CFX 1000 thermal cycler (Biorad Laboratories) and analyzed with the Bio-Rad CFX manager software V2.0 (Biorad Laboratories). All reactions were performed according to manufacturer's conditions, and consisted of 0.6 µL H_2O, 0.2 µL of forward and reverse primer each, 5 µL of Ssofast 2X master mix (Bio-Rad Laboratories) and 4µL of cDNA diluted to 40X.In all cases, optimum annealing temperature, primer specificity and amplification efficiency were determined using a PCR temperature gradient, a single melt-curve peak as well as gel electrophoresis and a serial dilution (Taylor *et al.* 2010), respectively, using a cDNA mixture containing equal amounts of cDNA extracted from each isolates alone condition. For all primers used, a 57^{o}C annealing temperature provided robust PCR amplifications of the desired target sequence, and upper and lower limit of reaction efficiency thresholds with the serial dilutions of 90% and 110% with a minimum R^2 value of 0.98 were obtained, suggesting optimal qPCR conditions (Taylor *et al.* 2010). Amplifications were performed using the following conditions: initial denaturing of 95^{o}C for 3min followed by 40 cycles of 95^{o}C for 10s and 57^{o}C for 6s and a final melt curve from 65^{o}C to 95^{o}C, with 0.5^{o}C increments, holding at each step for 5 s, with no template controls included for every target in each run.

Target gene expression was determined by Cq (cycle of quantification) values, using the single baseline threshold provided with the CFX software V2.0 (Biorad Laboratories) (Reller *et al.* 2011; Schneider *et al.* 2012). Inter-run calibrations within the CFX manager software were implemented following manufacturers recommendations to compare experimental conditions which were measured on separate QPCR runs and used in order to normalize for inter-plate variability (Bustin 2005). In this case, the software adjusts Cq values of all samples for each target gene between runs using the pairwise difference of each target gene amplified from a common sampledefined by the software, which was present on both runs.

To calculate normalized gene expression of all samples measured with all target genes, the delta-delta Ct method (Schmittgen and Livak 2008) implemented in the CFX manager

Table 2. List of primers used for PCR and RT-QPCR methods in this study

Specific Primers		
HMG #	Forward Primer	Reverse Primer
1	CGGCTAACACCACATAACTTCTACCC	TTTGGTATCTTCCGGTTCGTTTTGCC
2	TCCTAACGCTTTTATAATTTATCGTAAGG	GAGAAAAAGATTTTGGTAACATTTCCCAATG
3	TGCACCTACAAAACCACAGAATGC	AAAAGTCTTAAGTTCGGTTTCATAGG
4	CCTTTCCCTCCAACACTCATTACCG	GCGTAAGCATCTTTGACTTCTTTGG
5	TCAAAATGTTGGCCACTACTGAAAAT	AATCTTCGATGAGAAAAGACGACGTA
6	ATAATCCGATCACTTTCAGTTCGG	GTGTCTTTCTGCGGCCAATCTTGC
7	TATTATACCACGACGCCAAAATGCGT	GGTGTAGTTGGTGGGGTTATCGGTAT
8	ACCACCATTTCCACCTTTAATCG	CATCGCTCAATGATAAATCTGGACT
9	CTCCGAATGCTTTCATACTTTATCG	TCTTCTTGGACGATACCTATATTCTGG
10	GAATCTCGAACCTGCTTTTGTACAGC	GCCGTCTTTTGATGTAATCCTTTGGC
11	TCAAAAGGATTTAACCTCCCAATGCAA	TGGTTTGCTCTGAAAGTTCTTTATGTTG
12	GATACTCCTCCAAGACCTTTAAATTC	CAATAAACGTAATTTGGATACAATTTCATATG
13	TTTAACAAGCGAACATTATGCACC	GCATTCCTCGCTGATTGTGCAAAG
14	ATAATGTTCCCGCTACTGCGGCTAAT	ACCCAGCTTGAGATGCATCAAATACG
15	TCCGTTCCCACCAACAATTACTGC	TTTAATTCAACTTCAACTTGTGCGG
16	GTCCAAATGCATTCTTCGTTTATCG	CTTTCTCCTTCCAATATCAAAATCGTC
17	TATACCATCTCGATCTCCTAACGC	GTTGTTTCTTGCGTGGCGATTGCG
18	GAATGCTTTCATTCTTTATAGGCGTGC	CGTAATTAGGATAAGCTTGCATATGTTC
19	ATCTTTTTACCCGGCAGCACACCC	GACTTATTTGAGAAAGAAACAAAATTCC
20	CAAATGCTTTTATCCTTTATAGAAGAGC	TAATTTGGATAAGTTTGCATATGTTCC
21	ATCAAAGTTAAAACACCTCCC	CTCCATAAACCTCCCTGTGCT
22	AAACGAATTTGTCACGTGCTTGCG	TTGCTCTCTTTCTTCTACATCCTGA
23	TTTCCCGAATTAGATAAGCAAACC	CTTAAGTGGTATATATTTGTAACCTTCG
24	ACTGACGCGAAAGCTTCCTCAACC	ACGGAAGAGGCACGATAATGTCCG
25	AGGGTTATTAATCGGATTTGGTCCGT	CATGATATTCACGAGCATCATCACAAAA
26	ATCTAAACAAGTCGGGGAAATGTGGC	TTTGCTGCATCAGCTAAACGTTGGAA
27	ATTTATTCCGAACCTTGGGGTCGTTC	ATTGACTTGTTGCGGTTGTTGTTGTG
28	AGTAAAGTTCCTAGACCACAAAATCCT	TTTTTTTGGTTGATAAACATAATCAGGATA
29	TCAGGGAAGGCATATTTTGAATGCCG	ATACCTTAGTTTCGCCATCACCACCA
30	TCAAAATGCATGGATTATTTATCGTAGAG	TTATAATTTCCATACCTTTCGACGTGC
31	GGGACTCCAAGACTTATAGAGAACTCG	CGCTTGTTCCGAGAAATCAGTAACGGG
32	TTCCCTCCCAAACCTTACGTTGAAGT	TCTTGGAGGATTCACAGGAACGGAAA
33	TGTTTCCACCTCCAATGTCTGCCG	TGACCCATTTGACCCATTTGACCC
34	TGGAGAATTGAGTCGTCCACC	CAGTTTCTTGTCTATAATTCTCGGAATG

35	TCCCATTTCCGCCAATTATTAAGGCA	GAAACGGCCGGTCTGAACTTTGTATT
36	CTTCTAATGCATTTATGATATATCGTAAGG	GACAGATAGGTTCAAGATTGTAAGG
37	GACTTAGATCAAGAGCTCCGAATC	CTTTCGTGAATTATTGCTAGACTACTAC
38	TGCTCTTGAAAGGATGGCCAAGAAAA	GGGAAGAGGAGGATGACATCGACAAC
39	CGGTTTATCAATAAAACATCCAAGAAGA	TGGAAGCGGTAAAGGAAGATTCG
40	ATTTATCGAGCTGCTTTAGTTAAAGAACTC	GAATTTAACATTAATGAAGAACGTGCG
41	AAGAGCACCAACATCAACCATCAACC	TCCGGCATAGGTCTGTTGCTATCATT
42	ACATTGAATTCGTCAAAACCTCCAC	TGGACATATCCAGTCCATTAATAACACC
43	CCAAATGCATTCATGATTTATAGAATGC	CAACTACGGATCTTTGTTGAACGG
44	GTCAAAACGCTTGGATACTTTATCGC	AATTTTTACCGTATTGTTCAATGTGACG
45	AAAATGGAAACCCACCGCCTCC	TTGGGTATTTCTCGCGATGTCTCC
46	GACTCCACATCAAAACCTTCCACGAC	CATGGCCGCTTGCTTTCTATGTTCTA
47	CGAAATATTGCGACATCAAGATTCAG	GTCGTATAGTTTATAAGTAGTACG
48	CCATTTATGATTTTCAGGACTGCC	TTCTTGCTTTCGGCCTATATTCATACC
49	TTTTCCGCCAATCATAAACGTGGAGG	CAAGGAGTCGTTATCAATGATTGGTTCA
50	TTCGTTACGAGCATGTGTATTTATCG	AAGCGATATTTGCATTCCAAGGTTCC
51	CCATTCATTCGACCTCCTTTTCCACC	AGTACCGTCCGCTCCATCTGATAAAT
52	CCCAAATTATCACAAGCTGAAATGG	TGTTGCTGATGTTGCTGATGTTGC
53	TTCGTAATATGAGACAAGGACAAG	GGGAAACAATCTTGCTGCGGAATTGA
54	GTTTTACCTTCCCAACTCACATTCC	GGCATCGATTCCCATGGTTTAGTCG
55	CAGAAATGCACACCACCTCGATAAAA	TTGGATTTCAGCCTCCGTTCTTTTGC
56	CCAAGATATAATCGATGGCCACTCCC	TCAGTTGGAGAGGTTGTCAAAGAAGC
57	TCATATGCCATGACGGATATTTCAGG	TCTTTGTCTGATTGTCGTGATTTCGC
58	TTTCCAATATTACTGCTCATGAGC	CCGGAAATCGCAGATGTAACAA
59	TCCAACTTGTTCTGTTACTTCTTCC	ACAGAATTTGATCTGCAATGACCC
60	CGGATGATAAACCTTTCGAAGATCCA	ATTCGCTTAAACGTTGAAATTTGGCA
61	CTCGTTTACAATTTCGGGCACCGTAT	GAATCTTCCTCCTCCTCAGGTAGCAC
62	CTGGCGGTATCTCGAACGCA	GGTGAGAGTGATGGGAGGAGTGA
63	ATGGCCGCAAAACGAAATGTACCTAA	TGATGATCAACAGAAACATTTGCAC
64	CTATGCAGATTACCAGGAGGCTAGGG	CCGTCGGAATAACAAAGGATTTCATGC
65	TCACCTCCTCCAATCCAACTTCCT	TCGTGGGCTCATCAGACGAT
66	CAACTAGCAATGTTAAACTGCGAACG	TTCCACGTGACTAATTGGGTATCTCG
67	TTCCTTACCCACAAACTTTAGATCCGA	TCATGAACCCATTACAAGGTTTACGACT
68	CTGGAAAGTGGTGGAGTTCTGAGTCT	TGAAGCAAAGAGAAAGTCGTCATCAA
69	CCGCTCTTTAAAGTACCATTTCCACCA	CCTCAGATAGGGAGAGTTTATGCGGG
70	ATCTCTAGCCCAGATTCACTTCCG	TGGAGGAGCTATTATCGCTGTTGG
71	GGAAATATGACATCTTCTTCGGTCG	AACCAAGCTCGTTTGTTGGTTTCG
72	AGAATCGGTGAGAATTCCTAATGCT	CCACGAGCTTGATTATGTAG
73	GTCCCAATGCTTTCTTCGTATATAGG	GAAATTATAATTCTTTCAACTTCACGAGC
74	CGTCGTGCTATTCCTCGTCC	GTGTTTTGGTTGATAAACATATTGAGGATA

75	TGTAAGGACTGTTAAAAATCCTCGAC	GCTTGTCTCTTTGGTTGATAAACATAAC
76	AATGATCCCTCGTACTCGCAAGCC	ACCTCCATTCCTTGTAGTCATATCC
Q-PCR Primers		
HMG6-A	GACGCCAAAATGCATGGATTCTT	ACCACTTCTTTCGGCTCATCG
HMG6-B	ACGACGCCAAAATGCATGGGTATTA	GCTGCTTCGTATACTTTTATTACCTCC
HMG 1	TTCCGCGACCTCCTAATGCTTTCATA	GGTTCGTTTTGCCACATTTTACCGAC
HMG 37	AAGTTCCTCTTGGTCAAAAGAGCC	GGTGGTGGTGGGGAGGAAAG
HMG65	AACACATGACCGGGATGAGAGTACAA	TGTGCACCACTTTGTTGGTTTACTACA
HMG 52	AAAGGAGGAGGAGGAAAAGTTCGTGA	AGCTTCTGTTGTAGGTTCAGTCAAAGA
HMG 22	ACCTTTTCCTCCTGTTATTGATCCGGT	AAGCATTAGGAGCTCTCGCAGGAATA
Tandem Region Primers		
Gap1	CCACGTCCTCCAAACGCGTATATGAT	TGCGAAACTGCGGGATCGACAG
Gap2	AAAAGTTATCTTTGCCAATCCGGGTCT	ACCGGTATATGTCGATGCGGACT
Gap3	GTGTGAAGATCAATTGCTCCCCTTCTC	ACTTCATCCGAACTGAAAGTGATCGG
Gap4	ACGAAGCAGCAGCAAGATTGGC	AGGGTGCGAAACTGCGGTGA
Gap5	GGGTGTAGACTCTGAAAGGATTGAACC	GTGTGCCGACTGTGCCGTGT
Gap6	TCATGCTAATCCGTAGTTTTCGCCATC	CCCTGTGGCAAGCCCTGTGG
Gap7	CGCAATTAAAGTCCGCACATCATCAA	TCTGCGGCCAATCTTGCTGCC

software. First, relative quantities of each sample for a particular target gene was determined using the following formula:

$$RQ = (0.01E) \cdot 2^{(Cq\ min\ -\ Cqsample)},$$

Where E is the reaction efficiency, 2 represents a doubling of fluorescence values per PCR cycle, Cq min is the lowest Cq value of all samples measured with a particular target gene, and Cq sample is the Cq value of the sample/ target gene combination of which the relative quantity is being determined.This scales all samples to the same baseline (i.e. the sample with the lowest Cq value for a target gene will be set at a relative quantity of 1 and all other samples measured with the same target gene will be greater than 1).

Then, normalized gene expression of a target gene for a sample is determined by the formula:

$$\text{Normalized gene expression} = RQ_{\text{target gene sample}} / RQ_{\text{Ef1}\alpha\text{ sample}}$$

Where $RQ_{\text{target gene sample}}$ is the relative quantity of a sample with a particular target gene and $RQ_{\text{ef1}\alpha\text{ sample}}$ is the relative quantity of the same sample with EF1α. In our study EF1α was used an internal control, because this housekeeping has been previously used for the same purpose in QPCR studies of other fungi (Fang and Bidochka 2006; Landi *et al.* 2012;van Aarle *et al.* 2007). This allows for normalization of the relative quantity for all target genes using one internal control gene that remains at a constant expression level across all treatment groups. In a single case, one target gene failed to amplify from a one biological replicate of an experimental condition, so the average Cq value of the remaining biological replicates measured with the same target gene were used to replace the missing value (Silberberg *et al.* 2009). To ensure stable expression levels of EF1α across treatment conditions, we also included the measurement of β-Tubulin expression which is another housekeeping gene that has been used as a control in previous studies of gene expression (Landi *et al.* 2012). In our case, the relative expression of β-Tubulin normalized to EF1α is expected to be stable across all samples analysed and experimental designs we performed in the present study.

StudentT-tests (two-tailed) assuming unequal variances for all comparisons were used to measure statistically significant changes in a target gene expression between growth conditions. Changes in expression were considered statistically significant at a $P \leq 0.05$(Deshmukh *et al.* 2006; Francesconi *et al.* 2006).

Conventional PCR procedures

Conventional PCR procedures were used to isolate orthologues of the 77 *MAT*-HMG genes identified across independent transcriptome datasets from several isolates of the *AMF R. irregularis*. In these cases,fragments were amplified in 50µL reaction volumes using 2X Econotaq (Lucigen) mastermix, 0.2µL of each primer pair (10mM), and 10ng of template DNA. PCR programs generally followed: 95^{o}C for 3 minutes, followed by 34 cycles of 94^{o}C for 30s, 55^{o}C for 30s, then 72^{o}C elongation temperature for 60s, and a final elongation for 5 minutes at 72^{o}C. PCR program annealing temperature and elongation time were adjusted for certain difficult amplifications. In particular, for some adenine-thymine rich regions, PCR elongation temperature was reduced to 60^{o}C. PCR amplicons were resolved on 1.5% agarose 1x tris-acetate-EDTA gels in all cases, and were subjected to direct Sanger sequencing. PCR products displaying double peaks in chromatograms were re-amplified and subjected to bacterial cloning (Agilent) followed by plasmid sequencing. Sequence chromatograms were manually trimmed and inspected for errors.The resulting sequences were aligned onto existing transcripts using MUSCLE (Edgar 2004), or using the "Map to reference" assembly function available within the Geneious software package (Biomatters).

Inverse PCR procedures

Genomic DNA (~100 ng) of isolate SwiC2 was digested with restriction enzymes, the digests were precipitated with sodium acetate and 100% ethanol, and used in ligations with T4 DNA ligase (20 µl, 4°C for 16 h). The ligation reactions were used directly as the templates for PCR. 11 restriction enzymes were used. ClaI, EcoRI, HindIII, KpnI, NdeI, PstI, XbaI and XhoI recognize 6-bp sites. Three enzymes, BfaI, HpaII and Sau3AI, recognize 4-bp sites. The restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA, USA). Two PCR conditions were employed with *Ex Taq* (Takara, Shiga, Japan) in an Eppendorf Mastercycler. The parameters were either 94°C 2 min; then 32 cycles of 94°C 20 s, 50°C 20 s, 60°C 4 min, or 94°C 2 min, then 32 cycles of 94°C 20 s, 50°C 20 s, 68°C 3 min. PCR reactions were resolved on 0.8 % agarose 1x Tris-acetate-EDTA gels. PCR products were purified from agarose gel slices, and directly sequenced. In cases where amplicons were faint, they were cloned into plasmid TOPO pCR2.1 (Invitrogen/Life Technologies, Grand Island, NY, USA). Independent plasmid clones were sequenced using the universal M13F and M13R primers, and internal primers. The sequence reads were assembled with Sequencher software version 4.8 (Gene Codes Corporation, Ann Arbor, MI, USA).

Crossing experiments

Two experimental designs, based on in-vitro culturing of AMF, have been used in the present study to test for the involvement of some *R. irregularis MAT*-HMG in the process partner recognition . The first type of in-vitro culture represents a negative control. In this case, one single isolate is grown in the presence of M-media, so the RNA isolated from cultures represents a collection of transcripts originating from the mycelial network of one single isolate; these were called “standalone” cultures. (See Figure 1A).

The second type of culture was designed to harvest hyphae originating from two interacting mycelial networks; these were called the “crossing” cultures (See Figure 1B). Crossing cultures consist of 150mm circular plates with sub-compartments created with 70mm circular plates. 70mm plates contained M-media with sugar while the area within the 150mm plate but outside the 70mm plate contained M-media without sugar to avoid extensive root proliferation outside the 70mm plate container. “Standalone” plates contained one 70mm plate while crossing plates contained two 70mm plates. Hyphal exit points (two in the crossing cultures, and six standalone plates) were created in the 70mm plates by heating 1cm wide tweezers and melting openings in the edge of the70mm dish down to the level of the M-media. M-media bridges were later produced using a pipette across all exit points.

In our crossing experiments, a total of three isolates have been used; namely the isolates SwiA4, SwiC2 and SwiB3 of *R. irregularis*. These were chosen because they originated from the same experimental field in Switzerland (Croll *et al.* 2008; Koch *et al.* 2006;Koch *et al.* 2004) but genetically diverge from each other; thus facilitating their genotyping. Each of these three isolateswere grown using the two above mentioned cultures using the following conditions:"standalone", "self-crossed" where two mycelial networks of the same isolate were present (referred to as A4-A4, C2-C2, or B3-B3), or “outcrossed” where two mycelial networks of different isolates were present (referred to as A4-C2, C2-

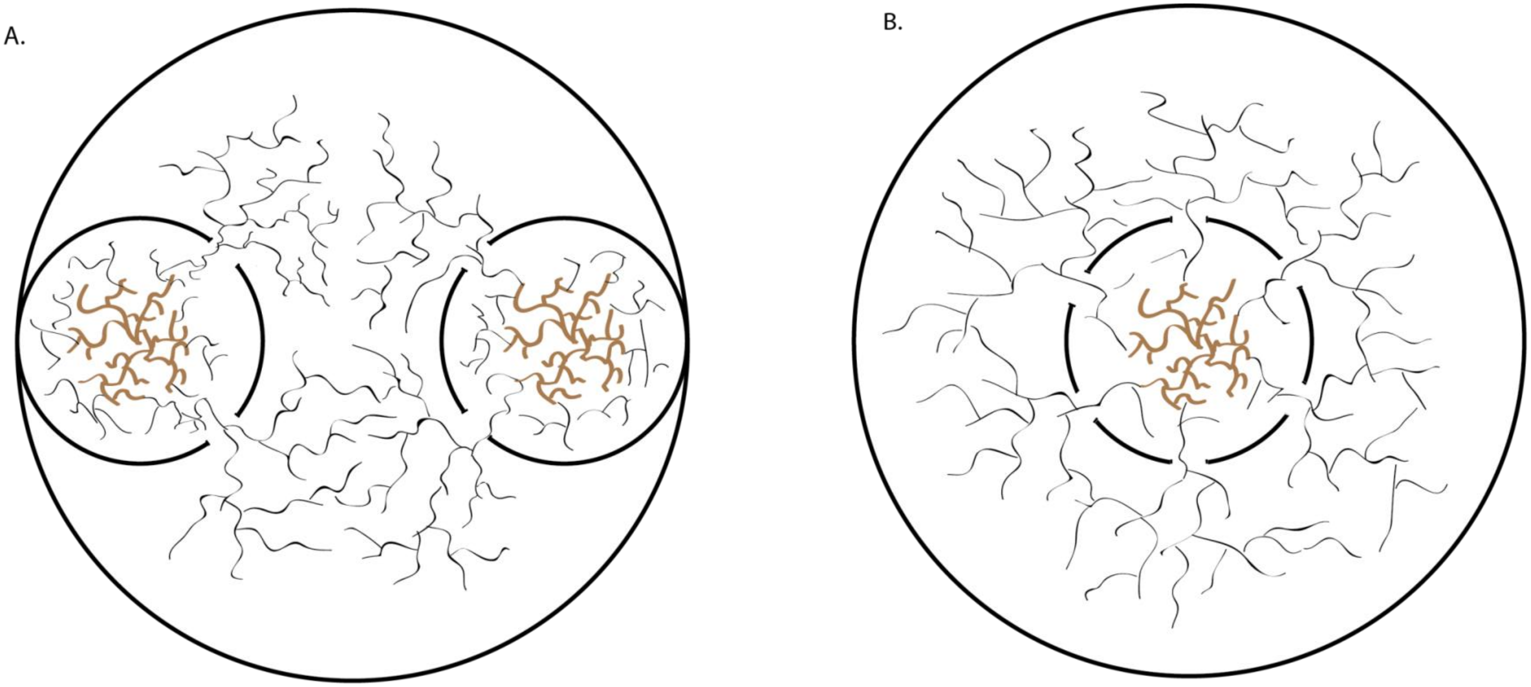


Figure 1. Drawing of the crossing(a) and standalone(b) cultures. Smaller circles within the larger circle are 70mm plates inside the 120mm plates. Gaps in small circles are exit points. Brown lines inside the 70mm plates are carrot roots. Finer black lines coming out of the carrot roots are hyphae.

B3, or A4-B3). Throughoutthe AMF culturing period, which always lasted a total of 32 days at 25^{o}C, carrot roots growing out of the 70mm dish were redirected back into the sub-compartment, while hyphae were allowed to grow through the exit points containing m-media bridges, and thus proliferate and interact with other hyphae in the same compartment. To obtain good RNA yields, we pooled two plates per each biological replicate.

Identification of recombination events

Recombination events were detected using RDP 4.13 which allows the application of several tools to detect the presence of recombination events (Martin *et al.* 2010). We chose the following recombination detection methods to detect recombinant sequences: RDP (Martin and Rybicki 2000), GENECONV (Padidam *et al.* 1999), Chimaerea (Posada and Crandall 1998), MaxChi (Smith 1992), BootScan (Salminen *et al.* 1995), SiScan (Gibbs *et al.* 2000), and 3seq (Boni *et al.* 2007). These methods detect recombination based on phylogenetic methods (BootScan, SiScan, RDP) or differences in nucleic acid positions (GeneConv, Chimaera, MaxChi, and 3Seq) along an alignment (Martin *et al.* 2010). All settings in RDP were left as default with the exception that sequences were set to linear (as opposed to circular).

Auto sequence masking was applied to alignments which removed highly similar sequences from the scan for recombination. This option keeps bonferroni p-value correction for multiple comparisons to a minimum, thus improving the recombination detection power (Martin *et al.* 2010). Since the program infers a potential recombinant and two potential parental sequences, we applied sequence masking to all sequences in an alignment but three sequences for each scan for recombination. We subjected putative recombinant sequences identified by at least five recombination detection method were used for further phylogenetic analysis. Previous studies have used between one and six tests showing evidence of recombination as a cutoff (Croll and Sanders 2009; Sentandreu *et al.* 2008), however we chose five as a conservative cutoff to screen out weak recombination signals.

When recombination was detected among isolates at a single *MAT*-HMG locus, the neighbour-network algorithm available in Splitstree v4.6 (Huson 1998) was used to test for reticulate branching patterns. This was carried out using *p*-distance and 1000 bootstrap replicates for branch support on a single HMG loci amplified from several individuals. Furthermore, we applied the Φ_w test for recombination in Splitstree using default settings as an additional measure of recombination (Bruen *et al.* 2006). Second, sequence alignments on either side of suggested recombination breakpoints were analysed independently by phylogenetics since different parts of a recombinant sequence will have different phylogenetic histories. For these alignments, we first selected the best-fit model of nucleic acid substitution among 88 models implemented in JmodelTest(Posada 2009) using the Akaikes information criterion corrected for small sample sizes (AICc). Maximum likelihood trees were constructed using PhyML 3.0(Guindon *et al.* 2010), with default settings with the exceptions of the starting tree topology search set to best of NNI and SPRand 100 bootstrap replicates for branch support.

Results

The Genome of Rhizophagus irregularis contains an unusually high number of MATA_HMG domains

Our searches for homologues of the *MAT*-HMGs across three available transcriptomes from *Rhizophagus irregularis* isolates and *R. diaphanous* resulted in the identification of a total of 76 gene candidates, all of which were found to harbour the motif defining the HMG domain of fungal mating type loci (i.e. MATA_HMG; CDD ID: cd01389). One additional HMG was also retrieved using inverse PCR in subsequent analysis. Importantly, all but one *MAT*-HMG gene (HMG10 which we failed to amplify from any of the isolates available in our laboratory), were found to be shared between the genomes of all AMF isolates analysed in this study following PCR and Sanger sequencing with specific primers (Table 2).

The total number of proteins with a *MAT*-HMG domain identified in *R. irregularis* far exceeds that found in any other fungal relative with a sequenced genome;with this specific gene family being at least 8 times more abundant in *R. irregularis* than it is in any fungus with a sequenced genome (Figure 2). Among the 76 *R. irregularis* genes harbouring a *MAT*-HMG domain, 3 were found to be most similar in sequence to the *SexM* (1) and *SexP* (2) genes that compose the mating type locus of several zygomycetes, while many others were found to be more closely related to *MAT*-HMGs from higher fungi (i.e. ascomycetes and basidiomycetes; *MAT 1-2-1, MAT 1-1-2, MAT Mc, Prf1*) (Table 3). A single transcript was found to contain two MATA_HMG domains (HMG74, Table 3). We also identified 8 potential pseudogenes (Table 3) among these AMF *MAT*-HMG domain containing

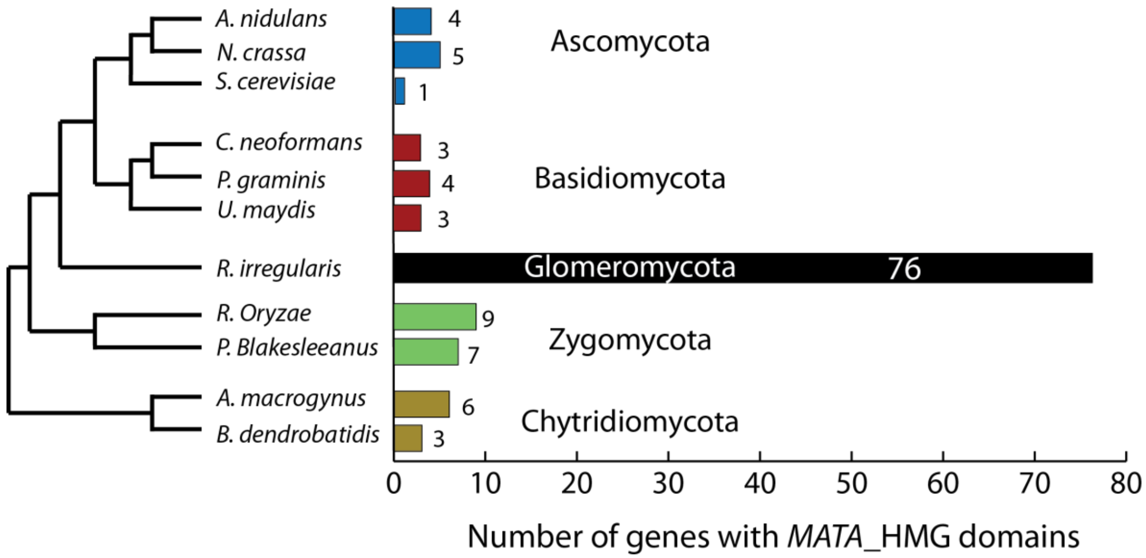


Figure 2. Number of *MAT*-HMG domain containing genes identified in our survey of the genome of the AMF *R. irregularis* and the genomes of representative species of the Ascomycota, Basidiomycota, Zygomycota, and Chytridiomycota.

Table 3: List predicted *MAT*-HMG domains found within *R. irregularis* and *R. diaphanous* isolates and the reciprocal blast first hit of each *MAT*-HMG domain containing query sequence. The query sequence accession along with the organism name, predicted domain name, functional domain annotation, e-value, and accession of the first blast hit retrieved from the NCBI nr databaseis listed for each *Rhizophagus spp. MAT*-HMG gene. The blue shaded row indicates a *MAT*-HMG which was identified from genomic scaffolds.

MAT-HMG #	**Query Accession**	**Organism**	**Annotation**	**Protein**	**E-Value**	**First blast hit Accession**
1	remain_C20901[1]	*Fusariumsacchari*	MATA_HMG	MAT1-2-1	0.003	BAE94382.1
2	BM959072.1[2]	*Xanthoriapolycarpa*	MATA_HMG	MAT1-2-1	0.001	CAI59768.2
3	remain_C19309[1]	*Colletotrichumhigginsianum*	MATA_HMG	HMG box protein	4.00E-12	CCF36618.1
4*	remain_C5083[1]	*Metarhiziumacridum*	MATA_HMG	HMG transcription factor	8.00E-05	EFY86728.1
5	KC785106[2]	*Grosmanniaclavigera*	MATA_HMG	MAT1-2 -1	0.004	EFX05114.1
6	KC785107[2]	*Verticilliumalboatrum*	MATA_HMG	predicted protein	4.00E-06	XP_003007798.1
7	KC517357[2]	*Melampsoralarici-populina*	MATA_HMG	hypothetical protein	9.00E-07	EGF99649.1
8	KC785103[2]	*Talaromycesmarneffei*	MATA_HMG	MAT1-2-1	0.003	ABC68485.1
9	GW085214.1[2]	*Schizosaccharomycespombe*	MATA_HMG	Mc 2	2.00E-10	NP_595867.1
10	KC785109[2]	*Penicilliumchrysogenum*	MATA_HMG	Hypothetical	4.00E-05	XP_002564591.1
11	KC785101[2]	*Erysiphenecator*	MATA_HMG	MAT1-2-1	0.079	AEB33764.1
12	GW085640.1[2]	*Diaporthe sp.*	MATA_HMG	mating type gene	8.00E-10	BAE93753.1
13	KC785098[2]	*Penicilliumchrysogenum*	MATA_HMG	mating type gene	3.00E-10	CCE33026.1
14*	GW086179.1[2]	*Mucormucedo*	MATA_HMG	SexM	4.00E-10	AFA26123.1
15	KC785099[2]	*Cryphonectriaparasitica*	MATA_HMG	MAT1-1-3	4.00E-04	AAK83344.1
16	GW088572.1[2]	*Trametesversicolor*	MATA_HMG	HMG box protein	5.00E-04	EIW55118.1
17	KC785097[2]	*Cladoniagalindezii*	MATA_HMG	MAT1-2	1.00E-04	AAT48651.1
18	GW088698.1[2]	*Fusariumoxysporum*	MATA_HMG	Mat-2 protein	4.00E-14	BAA28611.1

19	KC785108[2]	*Rhynchosporium secalis*	MATA_HMG	HMG box protein	3.00E-04	CAD62166.1
20	GW088846.1[2]	*Schizosaccharomycesjaponicus*	MATA_HMG	MatMc	1.00E-17	AFM85245.1
21	KC785100[2]	*Fibroporiaradiculosa*	MATA_HMG	predicted protein	0.002	CCM00606.1
22	KC785104[2]	*Trichodermaatroviride*	MATA_HMG	hypothetical protein	0.006	EHK50111.1
23	KC785105[2]	*Cercosporaapiicola*	MATA_HMG	MAT1-2	3.00E-10	ABB83710.1
24	KC785102[2]	*Rhynchosporium secalis*	MATA_HMG	HMG box protein	2.00E-10	CAD62166.1
25	KC785120[2]	*Fibroporiaradiculosa*	MATA_HMG	predicted protein	0.02	CCM01306.1
26	GW090102.1[2]	*Talaromycesmarneffei*	MATA_HMG	HMG box	6.00E-12	XP_002151220.1
27	KC785126[2]	*Rhynchosporium secalis*	MATA_HMG	HMG box protein	3.00E-04	CAD62166.1
28	GW093400.1[2]	*Trametesversicolor*	MATA_HMG	Hypothetical	2.00E-09	EIW63176.1
29*	KC814215[2]	*Talaromycesmarneffei*	MATA_HMG	MAT1-2-1	0.97	ABC68485.1
30	GW098009.1[2]	*Piriformosporaindica*	MATA_HMG	Hypothetical	5.00E-11	CCA67490.1
31	KC785117[2]	*Paracoccidioidesbrasiliensis*	MATA_HMG	MAT1-2	3.00E-05	AEI83491.1
32	GW098177.1[2]	*Penicilliumchrysogenum*	MATA_HMG	HMG box	0.006	XP_002564591.1
33	KC785121[2]	*Ustilagohordei*	MATA_HMG	Prf1	2.00E-11	CCF52951.1
34	GW103650.1[2]	*Talaromycesmarneffei*	MATA_HMG	MAT1-2-1	1.00E-04	XP_002152469.1
35	KC785119[2]	*Sordariafimicola*	MATA_HMG	Mat a-1	8.00E-04	CAB63226.1
36	GW103707.1[2]	*Rhizopusoryzae*	MATA_HMG	SexP	0.28	ADU04732.1
37	GW090024.1[2]	*Trametesversicolor*	MATA_HMG	hypothetical protein	1.8	EIW55066.1
38	GW111241.1[2]	*Schizosaccharomycespombe*	MATA_HMG	mc 2	5.00E-05	NP_595867.1
39	KC785113[2]	*Fibroporiaradiculosa*	MATA_HMG	Hypothetical	8.00E-05	CCM01306.1
40	GW112127.1[2]	*Gaeumannomycesgraminis*	MATA_HMG	Hypothetical	4.00E-04	EJT79613.1
41	KC785110[2]	*Aspergilluskawachii*	MATA_HMG	Hypothetical	2.4	GAA93066.1
42	GW115599.1[2]	*Zymoseptoriatritici*	MATA_HMG	Hypothetical	4.00E-04	XP_003855030.1
43	GW088880.1[2]	*Talaromycesmarneffei*	MATA_HMG	MAT1-2-1	0.008	ABC68485.1

44	GW118912.1[2]	*Cryphonectriaparasitica*	MATA_HMG	MAT1-1-3	1.00E-07	AAK83344.1
45	KC785115[2]	*Talaromycesstipitatus*	MATA_HMG	MAT1-2-1	2.00E-06	XP_002488738.1
46	GW120847.1[2]	*Trametesversicolor*	MATA_HMG	Hypothetical	1.00E-08	EIW52457.1
47	KC785125[2]	*Schizophyllum commune*	MATA_HMG	Hypothetical	3.00E-05	XP_003029891.1
48	GW122078.1[2]	*Piriformosporaindica*	MATA_HMG	Hypothetical	1.00E-04	CCA72393.1
49	KC785113[2]	*Syzygitesmegalocarpus*	MATA_HMG	SexP	0.093	AET35404.1
50	remain_C10144[1]	*Mycosphaerella populorum*	MATA_HMG	hypothetical protein	1	EMF17374.1
51	KC785111[2]	*Xanthoriapolycarpa*	MATA_HMG	MAT1-2-1	0.002	CAI59768.2
52	remain_C10514[1]	*Rhynchosporium secalis*	MATA_HMG	HMG box protein	2.00E-09	CAD62166.1
53	KC785128[2]	*Talaromycesmarneffei*	MATA_HMG	MAT1-2-1	3.00E-07	ABC68485.1
54	remain_C15333[1]	*Metarhiziumacridum*	MATA_HMG	HMG transcription factor	3.00E-07	EFY86728.1
55*	KC785127[2]	*Tremellamesenterica*	MATA_HMG	hypothetical	0.079	EIW72397.1
56	remain_C16561[1]	*Debaryomyceshansenii*	MATA_HMG	DEHA2E09460p	0.007	XP_459717.2
57*	remain_C20306[1]	*Metarhiziumanisopliae*	MATA_HMG	MAT1-1-3	0.005	BAE93596.1
58*	remain_C26978[1]	*Phomopsis sp.*	MATA_HMG	MAT1-2-1	0.058	AFP89369.1
59	remain_C2725[1]	*Pyrenophorateres*	MATA_HMG	Hypothetical	1.00E-16	XP_003297060.1
60	remain_C2727[1]	*Salpingoeca sp.*	MATA_HMG	Hypothetical	0.034	EGD75508.1
61	remain_C28337[1]	*Schizosaccharomycesjaponicus*	MATA_HMG	matMc	0.006	AFM85245.1
62	remain_C5967[1]	*Piriformosporaindica*	MATA_HMG	Hypothetical	1.00E-04	CCA72393.1
63[3]	remain_C6410[1]	*Acyrthosiphonpisum*	Sox_TCF	Hypothetical	1.00E-05	XP_003245920.1
64*	remain_C8213[1]	*Metarhiziumanisopliae*	MATA_HMG	MAT1-1-3	6.1	EFZ01123.1
65	remain_C2832[1]	*Colletotrichumhigginsianum*	MATA_HMG	HMG box protein	2.00E-04	CCF38267.1
66*	remain_C9612[1]	*Schizosaccharomycespombe*	MATA_HMG	matMc	0.75	AAB28876.1
67	KC785112[2]	*Colletotrichumhigginsianum*	MATA_HMG	HMG box protein	0.016	CCF38267.1
68	KC785123[2]	*Phycomycesblakesleeanus*	MATA_HMG	SexM	9.00E-06	ABX27909.1

69	KC785122[2]	*Verticilliumlongisporum*	MATA_HMG	MAT1-1-2	0.096	AEA29200.1
70	KC785124[2]	*Gaeumannomycesgraminis*	MATA_HMG	Hypothetical	0.002	EJT79613.1
71	KC785118[2]	*Verticillium dahlia*	MATA_HMG	Hypothetical	8.00E-06	EGY15843.1
72	GW104503.1[2]	*Hymenoscyphus pseudoalbidus*	MATA_HMG	MAT1-2-1	0.001	AFQ90566.1
73	GW083220.1[2]	*Colletotrichumhigginsianum*	MATA_HMG	HMG box protein	4.00E-04	CCF38267.1
74	remain_C3698[1]	*Baudoiniacompniacensis*	MATA_HMG	MAT1-2-1	2.00E-05	EMC98166.1
75	GW082685.1[2]	*Beauveriabassiana*	MATA_HMG	HMG box protein	1.00E-08	EJP65574.1
76	KC785116[2]	*Xanthoriapolycarpa*	MATA_HMG	MAT1-2-1	0.012	CAI59768.2

1. Transcripts can be accessed at the URL: http://mycor.nancy.inra.fr/IMGC/GlomusGenome/blast3.html

2. Transcripts can be accessed on the NCBI EST database

3. CDD search indicates this gene contains a MATA_HMG domain.

* Possible pseudogenes

transcripts, all of which were characterized by the presence of an early stop codon along the open reading frame surrounding the *MAT*-HMG blast hit.

Some MATA_HMG domains show substantial divergence between R. irregularis isolates from one population

To explore the possibility that idiomorphs of the 76 *MAT*-HMGs may exist in the genomes of AMF (i.e. the presence of two highly divergent alleles at one locus in different AMF individuals), we amplified the DNA region surrounding these *MAT*-HMG domains by PCR from three *R. irregularis* isolates (isolates SwiA4, SwiB3 and SwiC2), sequenced the resulting product using Sanger methodology, and explored the resulting chromatograms for the presence of allelic variation among isolates.All chromatograms showing the presence of double peaks were discarded from subsequent analyses to avoid the detection of false positives (i.e. allelic variation resulting from closely related paralogues, instead of orthologues).

Interestingly, the vast majority of AMF *MAT*-HMGs were found to be either monomorphic, or to vary only slightly at the sequence level between different members of the population. Specifically, only 35 AMF *MAT*-HMG geneswere found to be polymorphic at the amino-acid level between the three isolates (Figure 3), the most extreme being the HMG49 with an average amino-acid identity of 77% between pairsof isolates.Two *MAT*-HMGs not included in this comparison were HMG61, which was found to have a substantial 197bp deletion that was specific to isolate SwiA4, HMG10 which we failed to amplify from any isolate, and HMG55 which we failed to amplify from isolate SwiA4.

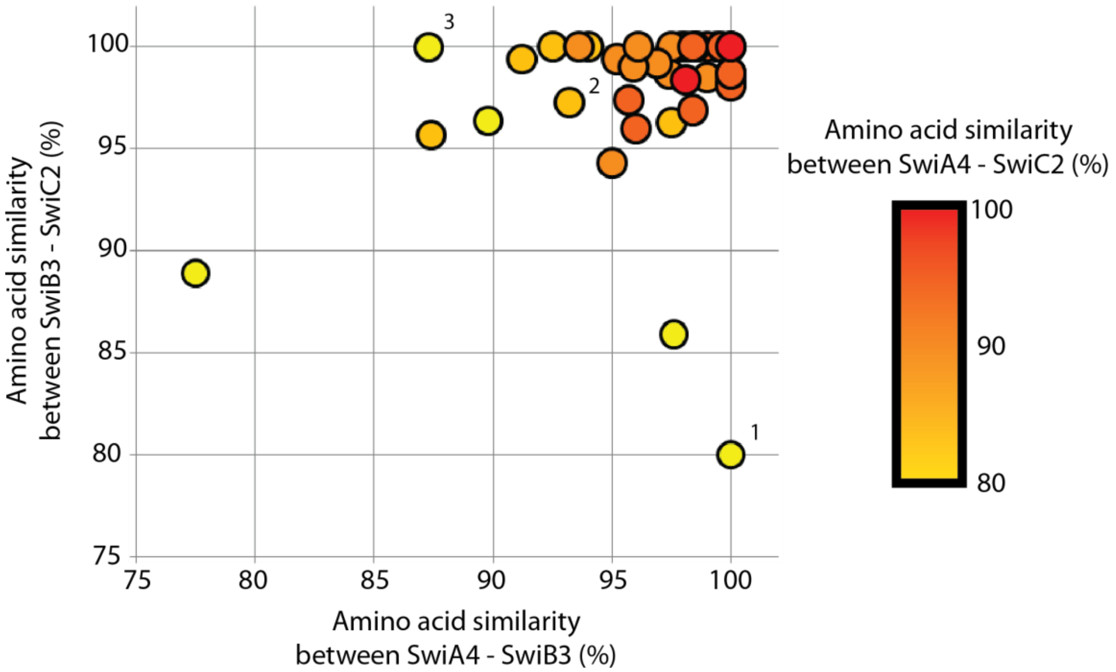


Figure 3. Pairwise amino acid similarity between *MAT*-HMG genes in isolates SwiA4, SwiB3 and SwiC2 of *R. irregularis*. Each colored circle represents a *MAT*-HMG gene and the color represents the % AA similarity between isolates SwiA4 and SwiC2. % AA similarity between isolates SwiA4 and SwiB3 is represented on the X-axis and between isolates SwiB3 and SwiC2 on the Y-axis. The numbers above some points represent variable *MAT*-HMG loci HMG6 (1), HMG1 (2), and HMG37 (3) which were analyzed by Q-RT-PCR.

Sequence variation was also investigated across additional isolates (n=17) for a number of *MAT*-HMGs of interest(HMG49, HMG37, HMG51, HMG6, HMG63, HMG61). These latter genes were chosen for their potential to represent *MAT* idiomorphs, as genes showed elevated sequence divergence and wererepeatedly found to be present in only one version of two divergent alleles among the three original isolatesinvestigated.Interestingly, specific allelic variation was sometimes maintained over extremely large geographical distances (i.e. *R. irregularis* isolates harvested from different continents harboured highly conserved polymorphisms); particularly in HMG49, HMG51, and HMG37, HMG61(Figure 4-7).

In four instances, potential alleles were found to be conserved across species boundaries where HMGs 49, 51 45, and 47, contained sequence polymorphisms which were conserved among strains of *R. irregularis* and one isolate of *R. diaphanous* (Figures 4, 5 and 8, 9).We also found isolates SwiC2 and Swi141 harboured potential isolate-specific pseudogenes the HMG61 locus resulting from one SNP (figure 7).

Origin and Nature of sequence diversity among homologues of the MAT-HMGs in the AMF

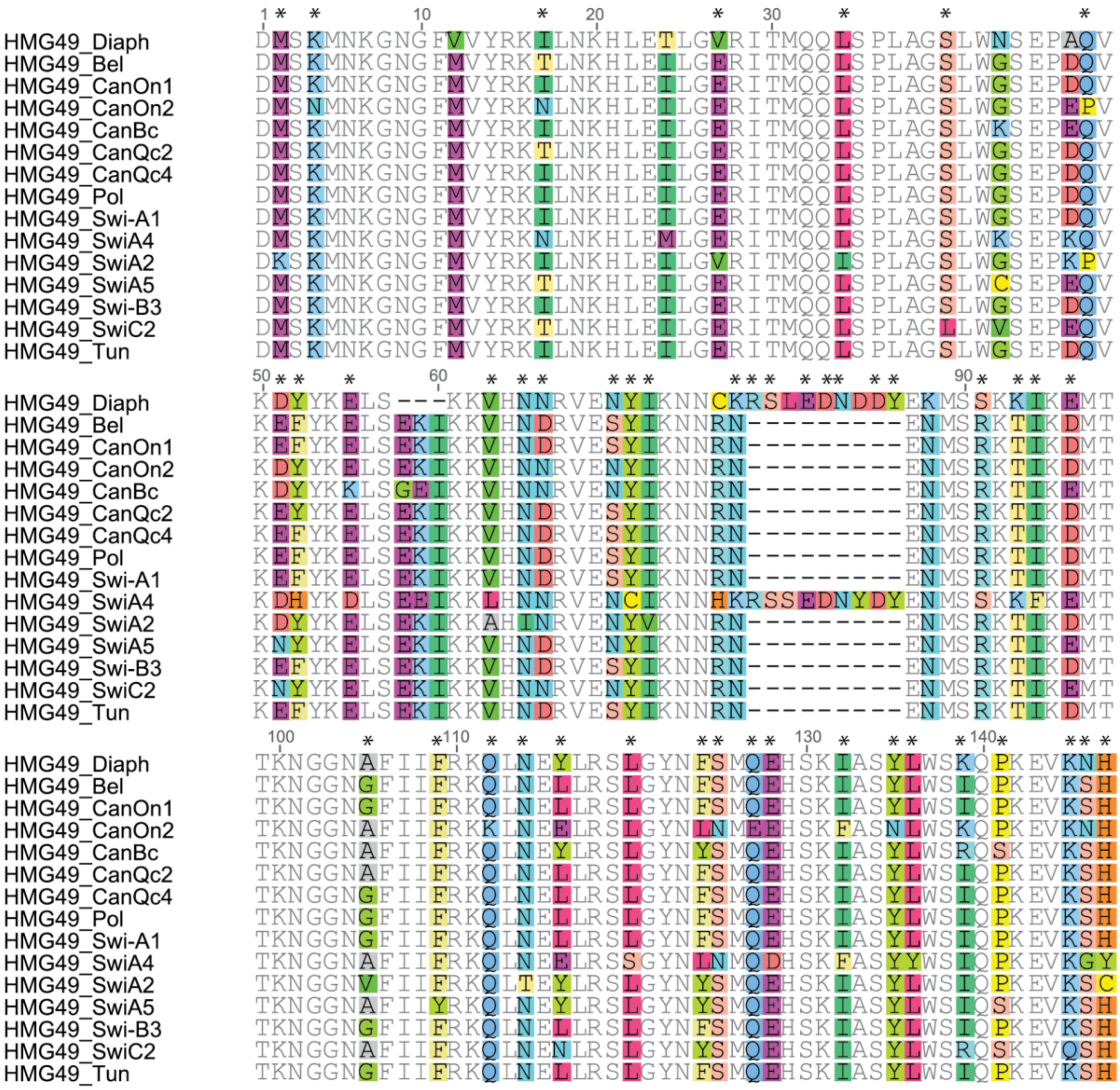
Detecting inter and intra-isolate recombination events in AMF

Previous studies have identified the presence of recombination across a number of AMF loci using various statistical tests which include phylogenetic and nucleotide differences based analysis (Croll and Sanders 2009; den Bakker *et al.* 2010; Vandenkoornhuyse *et al.* 2001).Given the number of homologous sequences (i.e. *MAT*-HMG genes) present in our dataset, a unique opportunity exists to further test for the presence of homologous recombination in AMF.

Aligning the nucleotide sequences of *MAT*-HMGs revealed the presence of a molecular diversity that could potentially result from recombination events within the genomes of AMF. Specifically, a number of genes were found to harbour sequences sharing strong similarities to partial sequences of other *MAT*- HMGs present elsewhere in the genome (Figure 10) and a single *MAT*-HMG displayed alterations in nucleotide similarity between isolates along an alignment (Figure 4). To determine if recombination was a primary mechanism driving the evolution of some *MAT*-HMGs and others, we applied seven detection methods implemented in RDP (Martin *et al.* 2010) which use a variety of phylogenetic and nucleotide-difference based tests to detect the presence of recombination breakpoints in aligned sequences. Our searches resulted in the detection of recombination involving a total of five *MAT*-HMG loci, and the presence of such events were supported by five or more of the recombination detection methods implemented in the RDP program (Table 4).

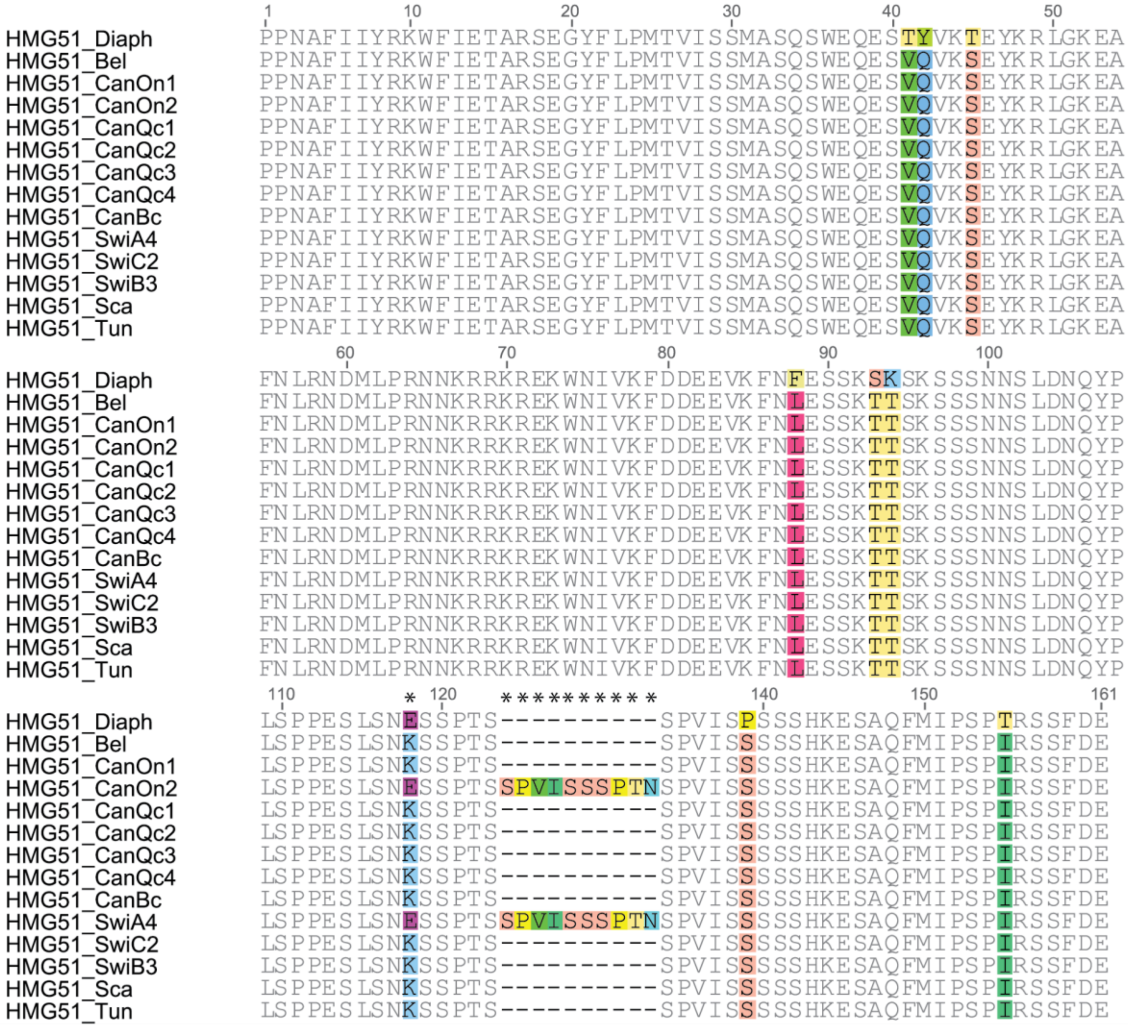
Two of these recombination events were suggested to have taken place within single isolates (i.e. homologous recombination between non-orthologous genes within one strain). Specifically, a portion of the HMG6 detected in isolate SwiC2 which was found to contain a 139bp fragment that is identical to HMG7from the same isolate. A second recombination event was suggested to have taken place between HMG 1 and HMG61 of isolates SwiC2 and SwiB3.

Potential for inter-isolate recombination was found in only one case, the *MAT*-HMG 49 (Table 4). In this case, a total of 8 potential recombination events could be identified, and these were found to have resulted in shuffling of genomic regions among several isolates of



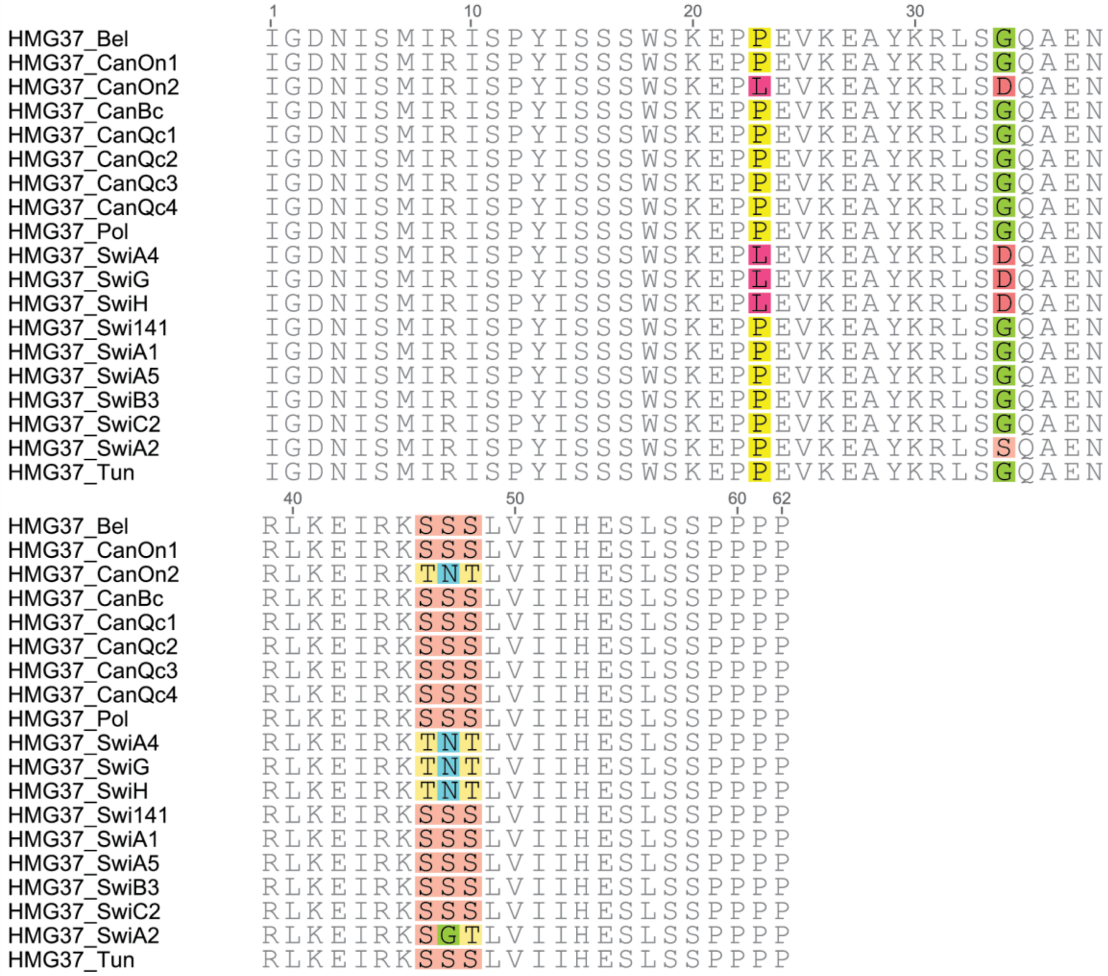
```
HMG49_Diaph    DMSKMNKGNGFVVYRKILNKHLETLGVRITMQQLSPLAGSLWNSEPAQV
HMG49_Bel      DMSKMNKGNGFMVYRKTLNKHLEILGERITMQQLSPLAGSLWGSEPDQV
HMG49_CanOn1   DMSKMNKGNGFMVYRKILNKHLEILGERITMQQLSPLAGSLWGSEPDQV
HMG49_CanOn2   DMSNMNKGNGFMVYRKNLNKHLEILGERITMQQLSPLAGSLWGSEPEPV
HMG49_CanBc    DMSKMNKGNGFMVYRKILNKHLEILGERITMQQLSPLAGSLWKSEPEQV
HMG49_CanQc2   DMSKMNKGNGFMVYRKTLNKHLEILGERITMQQLSPLAGSLWGSEPDQV
HMG49_CanQc4   DMSKMNKGNGFMVYRKILNKHLEILGERITMQQLSPLAGSLWGSEPDQV
HMG49_Pol      DMSKMNKGNGFMVYRKILNKHLEILGERITMQQLSPLAGSLWGSEPDQV
HMG49_Swi-A1   DMSKMNKGNGFMVYRKILNKHLEILGERITMQQLSPLAGSLWGSEPDQV
HMG49_SwiA4    DMSKMNKGNGFMVYRKNLNKHLEMLGERITMQQLSPLAGSLWKSEPKQV
HMG49_SwiA2    DKSKMNKGNGFMVYRKILNKHLEILGVRITMQQISPLAGSLWGSEPKPV
HMG49_SwiA5    DMSKMNKGNGFMVYRKTLNKHLEILGERITMQQLSPLAGSLWCSEPEQV
HMG49_Swi-B3   DMSKMNKGNGFMVYRKILNKHLEILGERITMQQLSPLAGSLWGSEPDQV
HMG49_SwiC2    DMSKMNKGNGFMVYRKTLNKHLEILGERITMQQLSPLAGLLWVSEPEQV
HMG49_Tun      DMSKMNKGNGFMVYRKILNKHLEILGERITMQQLSPLAGSLWGSEPDQV

HMG49_Diaph    KDYYKELS---KKVHNNRVENYIKNNCKRSLEDNDDYEKMSSKKIKEMT
HMG49_Bel      KEFYKELSEKIKKVHNDRVESYIKNNRN---------ENMSRKTIKDMT
HMG49_CanOn1   KEFYKELSEKIKKVHNDRVESYIKNNRN---------ENMSRKTIKDMT
HMG49_CanOn2   KDYYKELSEKIKKVHNNRVENYIKNNRN---------ENMSRKTIKDMT
HMG49_CanBc    KDYYKKLSGEIKKVHNNRVENYIKNNRN---------ENMSRKTIKEMT
HMG49_CanQc2   KEYYKELSEKIKKVHNDRVESYIKNNRN---------ENMSRKTIKDMT
HMG49_CanQc4   KEFYKELSEKIKKVHNDRVESYIKNNRN---------ENMSRKTIKDMT
HMG49_Pol      KEFYKELSEKIKKVHNDRVESYIKNNRN---------ENMSRKTIKDMT
HMG49_Swi-A1   KEFYKELSEKIKKVHNDRVESYIKNNRN---------ENMSRKTIKDMT
HMG49_SwiA4    KDHYKDLSEEIKKLHNNRVENCIKNNHKRSSEDNYDYENMSSKKFKEMT
HMG49_SwiA2    KDYYKELSEKIKKAHINRVENYVKNNRN---------ENMSRKTIKDMT
HMG49_SwiA5    KNYYKELSEKIKKVHNDRVENYIKNNRN---------ENMSRKTIKEMT
HMG49_Swi-B3   KEFYKELSEKIKKVHNDRVESYIKNNRN---------ENMSRKTIKDMT
HMG49_SwiC2    KNYYKELSEKIKKVHNNRVENYIKNNRN---------ENMSRKTIKEMT
HMG49_Tun      KEFYKELSEKIKKVHNDRVESYIKNNRN---------ENMSRKTIKDMT

HMG49_Diaph    TKNGGNAFIIFRKQLNEYLRSLGYNFSMQEHSKIASYLWSKQPKEVKNH
HMG49_Bel      TKNGGNGFIIFRKQLNELLRSLGYNFSMQEHSKIASYLWSIQPKEVKSH
HMG49_CanOn1   TKNGGNGFIIFRKQLNELLRSLGYNFSMQEHSKIASYLWSIQPKEVKSH
HMG49_CanOn2   TKNGGNAFIIFRKKLNEELRSLGYNLNMEEHSKFASNLWSKQPKEVKNH
HMG49_CanBc    TKNGGNAFIIFRKQLNEYLRSLGYNYSMQEHSKIASYLWSRQSKEVKSH
HMG49_CanQc2   TKNGGNAFIIFRKQLNELLRSLGYNFSMQEHSKIASYLWSIQPKEVKSH
HMG49_CanQc4   TKNGGNGFIIFRKQLNELLRSLGYNFSMQEHSKIASYLWSIQPKEVKSH
HMG49_Pol      TKNGGNGFIIFRKQLNELLRSLGYNFSMQEHSKIASYLWSIQPKEVKSH
HMG49_Swi-A1   TKNGGNGFIIFRKQLNELLRSLGYNFSMQEHSKIASYLWSIQPKEVKSH
HMG49_SwiA4    TKNGGNAFIIFRKQLNEELRSSGYNLNMQDHSKFASYYWSIQPKEVKGY
HMG49_SwiA2    TKNGGNVFIIFRKQLTEYLRSLGYNYSMQEHSKIASYLWSIQPKEVKSC
HMG49_SwiA5    TKNGGNAFIIYRKQLNEYLRSLGYNYSMQEHSKIASYLWSIQSKEVKSH
HMG49_Swi-B3   TKNGGNGFIIFRKQLNELLRSLGYNFSMQEHSKIASYLWSIQPKEVKSH
HMG49_SwiC2    TKNGGNAFIIFRKQLNENLRSLGYNYSMQEHSKIASYLWSRQSKEVQSH
HMG49_Tun      TKNGGNGFIIFRKQLNELLRSLGYNFSMQEHSKIASYLWSIQPKEVKSH
```

Figure 4. Amino-acid alignment of HMG49 showing the conservation polymorphisms in isolates from broad geographic locations and trans-species polymorphisms between an isolate of *R.diaphanous* and isolates of *R.irregularis* at amino-acid positions indicated by the asterisks. Note the alternating amino acid identities between isolates along the alignment which is a sign of potential recombination. Isolate codes can be found in table 1. Polymorphic amino acid positions along the alignment are highlighted and each amino acid is assigned a unique colour.



```
                1        10        20        30        40        50
HMG51_Diaph     PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESTYVKTEYKRLGKEA
HMG51_Bel       PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_CanOn1    PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_CanOn2    PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_CanQc1    PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_CanQc2    PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_CanQc3    PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_CanQc4    PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_CanBc     PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_SwiA4     PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_SwiC2     PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_SwiB3     PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_Sca       PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
HMG51_Tun       PPNAFIIYRKWFIETARSEGYFLPMTVISSMASQSWEQESVQVKSEYKRLGKEA
                     60        70        80        90        100
HMG51_Diaph     FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNFESSKSKSKSSSNNSLDNQYP
HMG51_Bel       FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_CanOn1    FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_CanOn2    FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_CanQc1    FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_CanQc2    FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_CanQc3    FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_CanQc4    FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_CanBc     FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_SwiA4     FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_SwiC2     FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_SwiB3     FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_Sca       FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
HMG51_Tun       FNLRNDMLPRNNKRRKREKWNIVKFDDEEVKFNLESSKTTSKSSSNNSLDNQYP
                 110     * 120 **********      140       150      161
HMG51_Diaph     LSPPESLSNESSPTS----------SPVISPSSSHKESAQFMIPSPTRSSFDE
HMG51_Bel       LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_CanOn1    LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_CanOn2    LSPPESLSNESSPTSSPVISSSPTNSPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_CanQc1    LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_CanQc2    LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_CanQc3    LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_CanQc4    LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_CanBc     LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_SwiA4     LSPPESLSNESSPTSSPVISSSPTNSPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_SwiC2     LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_SwiB3     LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_Sca       LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
HMG51_Tun       LSPPESLSNKSSPTS----------SPVISSSSSHKESAQFMIPSPIRSSFDE
```

Figure 5. Amino-acid alignment of HMG51 showing the conservation polymorphisms in isolates from broad geographic locations and trans-species polymorphisms between an isolate of *R.diaphanous* and isolates of *R.irregularis* at amino-acid positions indicated by the asterisks. Isolate codes can be found in table 1. Polymorphic amino acid positions along the alignment are highlighted and each amino acid is assigned a unique colour.



```
                    1         10         20          30
HMG37_Bel           IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_CanOn1        IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_CanOn2        IGDNISMIRISPYISSSWSKEPLEVKEAYKRLSDQAEN
HMG37_CanBc         IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_CanQc1        IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_CanQc2        IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_CanQc3        IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_CanQc4        IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_Pol           IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_SwiA4         IGDNISMIRISPYISSSWSKEPLEVKEAYKRLSDQAEN
HMG37_SwiG          IGDNISMIRISPYISSSWSKEPLEVKEAYKRLSDQAEN
HMG37_SwiH          IGDNISMIRISPYISSSWSKEPLEVKEAYKRLSDQAEN
HMG37_Swi141        IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_SwiA1         IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_SwiA5         IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_SwiB3         IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_SwiC2         IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN
HMG37_SwiA2         IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSSQAEN
HMG37_Tun           IGDNISMIRISPYISSSWSKEPPEVKEAYKRLSGQAEN

                     40        50        60 62
HMG37_Bel           RLKEIRKSSSLVIIHESLSSPPPP
HMG37_CanOn1        RLKEIRKSSSLVIIHESLSSPPPP
HMG37_CanOn2        RLKEIRKTNTLVIIHESLSSPPPP
HMG37_CanBc         RLKEIRKSSSLVIIHESLSSPPPP
HMG37_CanQc1        RLKEIRKSSSLVIIHESLSSPPPP
HMG37_CanQc2        RLKEIRKSSSLVIIHESLSSPPPP
HMG37_CanQc3        RLKEIRKSSSLVIIHESLSSPPPP
HMG37_CanQc4        RLKEIRKSSSLVIIHESLSSPPPP
HMG37_Pol           RLKEIRKSSSLVIIHESLSSPPPP
HMG37_SwiA4         RLKEIRKTNTLVIIHESLSSPPPP
HMG37_SwiG          RLKEIRKTNTLVIIHESLSSPPPP
HMG37_SwiH          RLKEIRKTNTLVIIHESLSSPPPP
HMG37_Swi141        RLKEIRKSSSLVIIHESLSSPPPP
HMG37_SwiA1         RLKEIRKSSSLVIIHESLSSPPPP
HMG37_SwiA5         RLKEIRKSSSLVIIHESLSSPPPP
HMG37_SwiB3         RLKEIRKSSSLVIIHESLSSPPPP
HMG37_SwiC2         RLKEIRKSSSLVIIHESLSSPPPP
HMG37_SwiA2         RLKEIRKSGTLVIIHESLSSPPPP
HMG37_Tun           RLKEIRKSSSLVIIHESLSSPPPP
```

Figure 6. Amino acid alignment of HMG37 showing the conservation of polymorphisms among isolates of *R.irregularis* from a broad geographic origin. Isolate codes can be found in table 1. Polymorphic amino acid positions along the alignment are highlighted and each amino acid is assigned a unique colour.

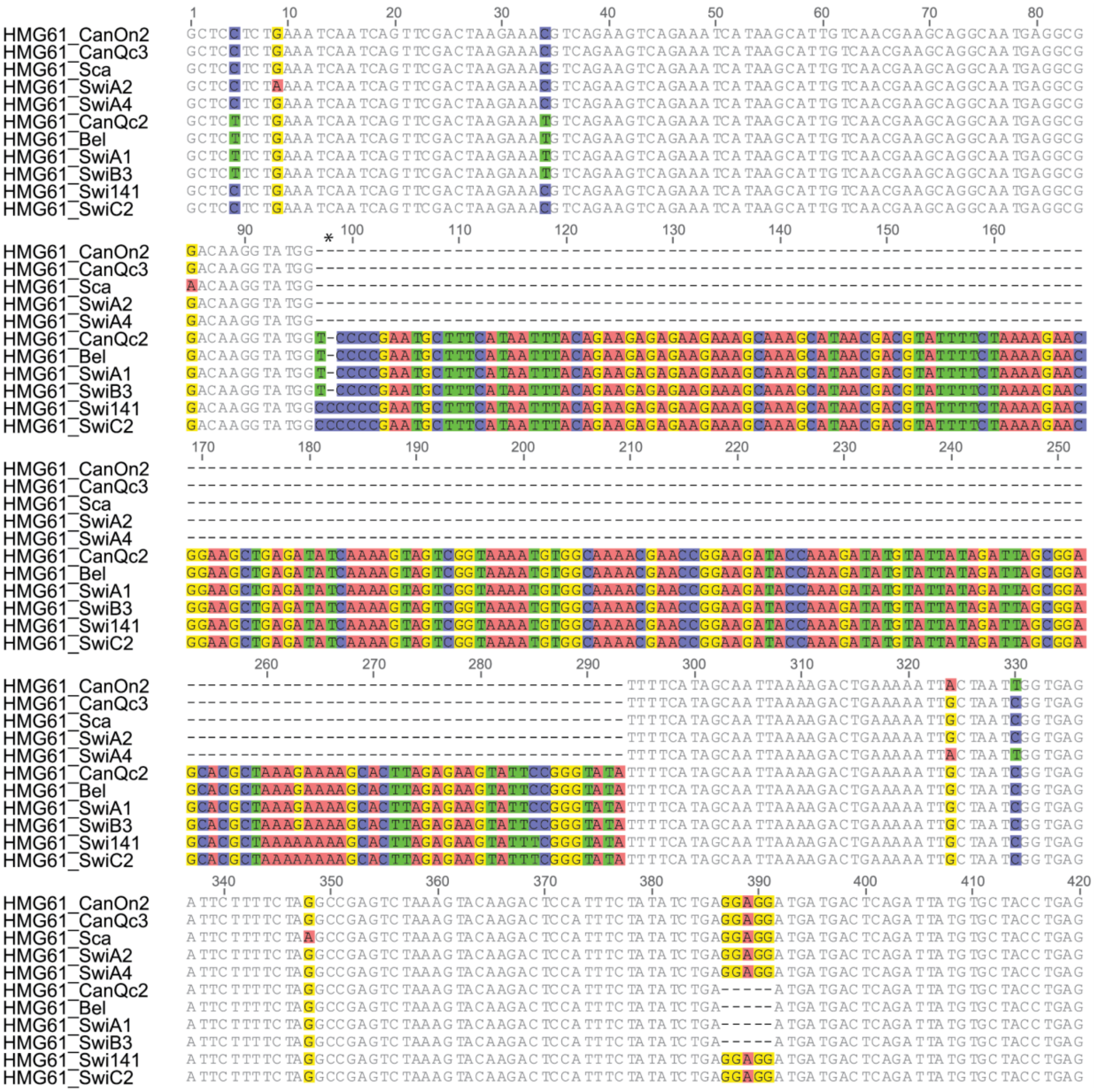
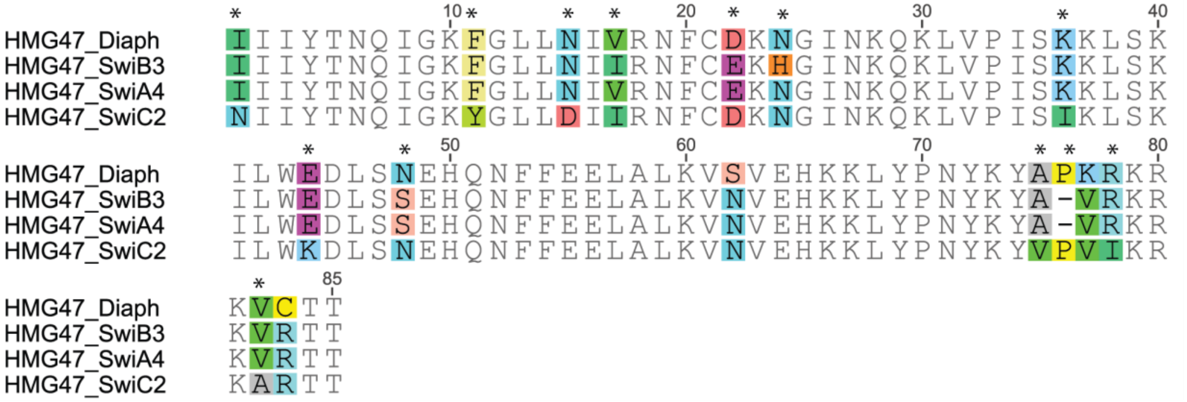


Figure 7. Polymorphisms conserved across isolates from broad geographic regions at HMG61. Nucleic acids are shown rather than amino-acids because of a frameshift mutation specific to isolate SwiC2 and Swi141 at nucleotide position 98 indicated by the asterisk. Isolate codes can be found in table 1. Polymorphic nucleotide positions along the alignment are highlighted and each amino acid is assigned a unique colour.

HMG45_Diaph LPCPRSADELQMPEKGTVPREMNSFIVFRRQLAHVAKLHKT
HMG45_SwiC2 LPCPRSADELQMPEKGTVPREMNSFIIFRRQLAHVAKLHKT
HMG45_SwiA4 LPCPRSADELQMPEKGTVPREMNSFIVFRRQLAHVAKLHKT
HMG45_SwiB3 LPCPRSADELQMPEKGTVPREMNSFIVFRRQLAHVAKLHKT

HMG45_Diaph SDDGRVISTAASFIWNGASKNEKKSYSLIAQELKRRHREK
HMG45_SwiC2 SDDGRVISTAASFIWNGASKNEKMSYSLIAQELKRRHREK
HMG45_SwiA4 SDDGRVISTAASFIWNGASKSEKRSYTLIAEELKRRHREK
HMG45_SwiB3 SDDGRVISTAASFIWNGASKNEKRSYSLIAEELKRRHREK

Figure 8. Amino-acid alignment of HMG45 showing trans-species polymorphisms between an isolate of *R.diaphanous* and isolates of *R.irregularis* at amino-acid positions indicated by the asterisks. Isolate codes can be found in table 1. Polymorphic amino acid positions along the alignment are highlighted and each amino acid is assigned a unique colour.



HMG47_Diaph IIIYTNQIGKFGLLNIVRNFCDKNGINKQKLVPISKKLSK
HMG47_SwiB3 IIIYTNQIGKFGLLNIIRNFCEKHGINKQKLVPISKKLSK
HMG47_SwiA4 IIIYTNQIGKFGLLNIVRNFCEKNGINKQKLVPISKKLSK
HMG47_SwiC2 NIIYTNQIGKYGLLDIIRNFCDKNGINKQKLVPISIKLSK

HMG47_Diaph ILWEDLSNEHQNFFEELALKVSVEHKKLYPNYKYAPKRKR
HMG47_SwiB3 ILWEDLSSEHQNFFEELALKVNVEHKKLYPNYKYA-VRKR
HMG47_SwiA4 ILWEDLSSEHQNFFEELALKVNVEHKKLYPNYKYA-VRKR
HMG47_SwiC2 ILWKDLSNEHQNFFEELALKVNVEHKKLYPNYKYVPVIKR

HMG47_Diaph KVCTT
HMG47_SwiB3 KVRTT
HMG47_SwiA4 KVRTT
HMG47_SwiC2 KARTT

Figure 9. Amino-acid alignment of HMG47 showing trans-species polymorphisms between an isolate of *R.diaphanous* and isolates of *R.irregularis* at amino-acid positions indicated by the asterisks. Isolate codes can be found in table 1. Polymorphic amino acid positions along the alignment are highlighted and each amino acid is assigned a unique colour.

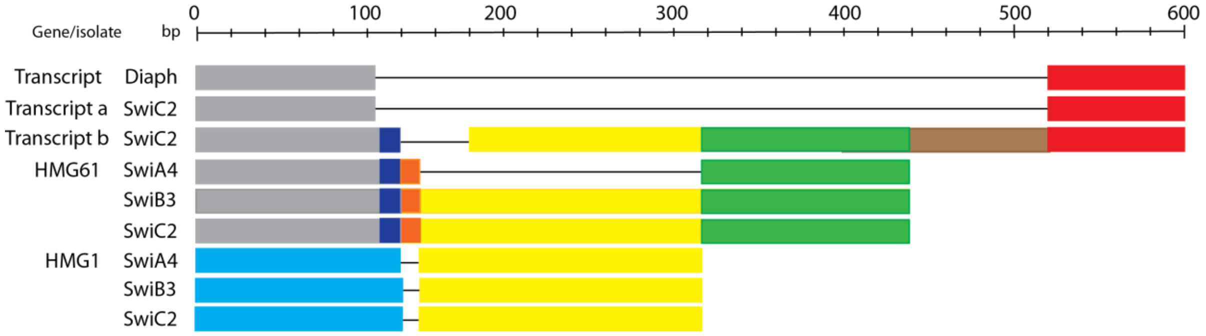


Figure 10. An alignment of four transcripts, two of which with sequences from three isolates of *R.irregularis*. Colored boxes represent sequences which are highly similar (>90%) and do not resemble the sequence represented by any other colored box. Grey and blue boxes are have an approximate 40% nucleotide identity. Lines on the alignment are gaps. The scale above the alignment represent nucleotide position.

Interisolate recombination

Recombinant HMG #	$49^{2,4}$	$49^{1,4}$	$49^{3,4}$	$49^{3,4}$	$49^{1,4}$	$49^{1,4}$	$49^{2,4}$	$49^{3,4}$
First breakpoint	101*	137	66*	81*	29*	345*	152	316*
Second Breakpoint	301*	359	324*	278	311*	438*	347*	440*
Recombinant sequence	CanOn2	CanBC	CanOn2	CanOn2	CanOn2	CanOn2	CanOn2	CanOn2
Major parent sequence	Swi A4	SwiA4	CanQc2*	Swi A1*	CanBc*	Swi A2	Swi C2*	SwiC2
Minor parent sequence	Swi A2	SwiC2	SwiA2	Swi A2	Swi A2	CanBc*	Swi A1	CanBc*
RDP	-	-	-	3.90E-02	-	-	3.20E-03	-
GeneConv	3.80E-02	0.043	0.0048	-	0.015	0.0089	1.10E-02	0.014
BootScan	1.40E-02	0.031	0.00095	4.20E-03	0.017	0.045	3.80E-02	0.024
MaxChi	1.70E-03	0.0028	0.0026	6.70E-03	0.0065	0.0029	5.60E-03	0.008
Chimaera	4.40E-02	0.007	0.044	6.70E-03	0.00014	0.0012	1.30E-03	0.025
SiScan	2.30E-07	2.4E-09	-	8.90E-03	0.0046	0.0041	2.40E-02	0.000005
3Seq	3.40E-02	0.011	0.0039	8.80E-04	0.00088	0.018	9.20E-03	-

Intraisolate recombination

Recombinant HMG #	7^{4}	1^{4}	1^{4}
First breakpoint	197*	150	147
Second Breakpoint	336	264*	264*
Recombinant sequence	HMG 6 - Swi C2	HMG 1 - Swi B3	HMG 1 - Swi C2
Major parent sequence	HMG 6 -Swi B3	HMG 1 - Swi A4	HMG 1 - Swi A4
Minor parent sequence	HMG 7 - Swi C2	HMG 61 - Swi B3	HMG 61 - Swi B3
RDP	-	-	-
GeneConv	4.70E-07	9.90E-06	1.30E-07
BootScan	1.30E-04	1.90E-07	3.10E-08
MaxChi	1.10E-09	4.60E-04	1.40E-03
Chimaera	3.60E-09	1.30E-03	1.30E-03
SiScan	2.30E-09	-	3.30E-09
3Seq	1.00E-09	6.60E-07	8.10E-07

Table 4. Results of RDP analysis of MAT-HMG seqeunces where recombination events were detected by five or more recombination detection tools. Warnings stated by the program are represented by numbers above each HMG in recombination event; 1- either parental sequence is a potential recombinant, 2- the major parent is the potential recombinant, 3- the minor parent is the potential recombinant. Asterisks next to recombnation breakpoints and isolates indicate an uncertain designation of a recombination breakpoint, minor parent or major parent by RDP.

R. irregularis. However, in this case it is important to note that suggested recombinants are prone to possible miss-identification.

Since recombination produces sequences with regions that have alternate evolutionary histories, we constructed phylogenetic trees using the sequence regions bounded by the recombination breakpoints estimated by the RDP analysis. Phylogenetic analysis significantly supported a conflicting evolutionary relationship between sequence regions of the recombinant HMG6 gene in isolate SwiC2 (Figure 11a). Unfortunately, the recombination event in between HMG1 and 77 was not significantly supported (data not shown). Since suggested recombination breakpoints clustered around the same region of the alignment on HMG49 we construct two trees from the 3' sequence region and the 5' sequence region of the alignment which produced two phylogenetic trees which significantly supported incongruent phylogenetic histories of the 5' and 3' sequence regions of HMG49 (Figure 11b).

The presence of recombination is expected to produce a reticulate branching pattern following phylogenetic reconstructions. For this reason weapplied the neighbour-network algorithmin Splitstree 4(Huson and Bryant 2006) to visualize any reticulate associations among isolates harbouring potentially recombinant of HMG 49, along with the $\Phi_w$test for recombination implemented within Splitstree. The combination of both analyses is considered a robust analysis that effectively distinguishes between recurrent mutation and recombination in an alignment (Bruen *et al.* 2006). In the case of HMG49, the neighbour network algorithm produced a phylogenetic network with reticulating branches between representatives of the various AMF isolates (Figure 12), and recombination within this reticulate network was significantly supported by the Φ_w test for recombination(p=0.00000074).

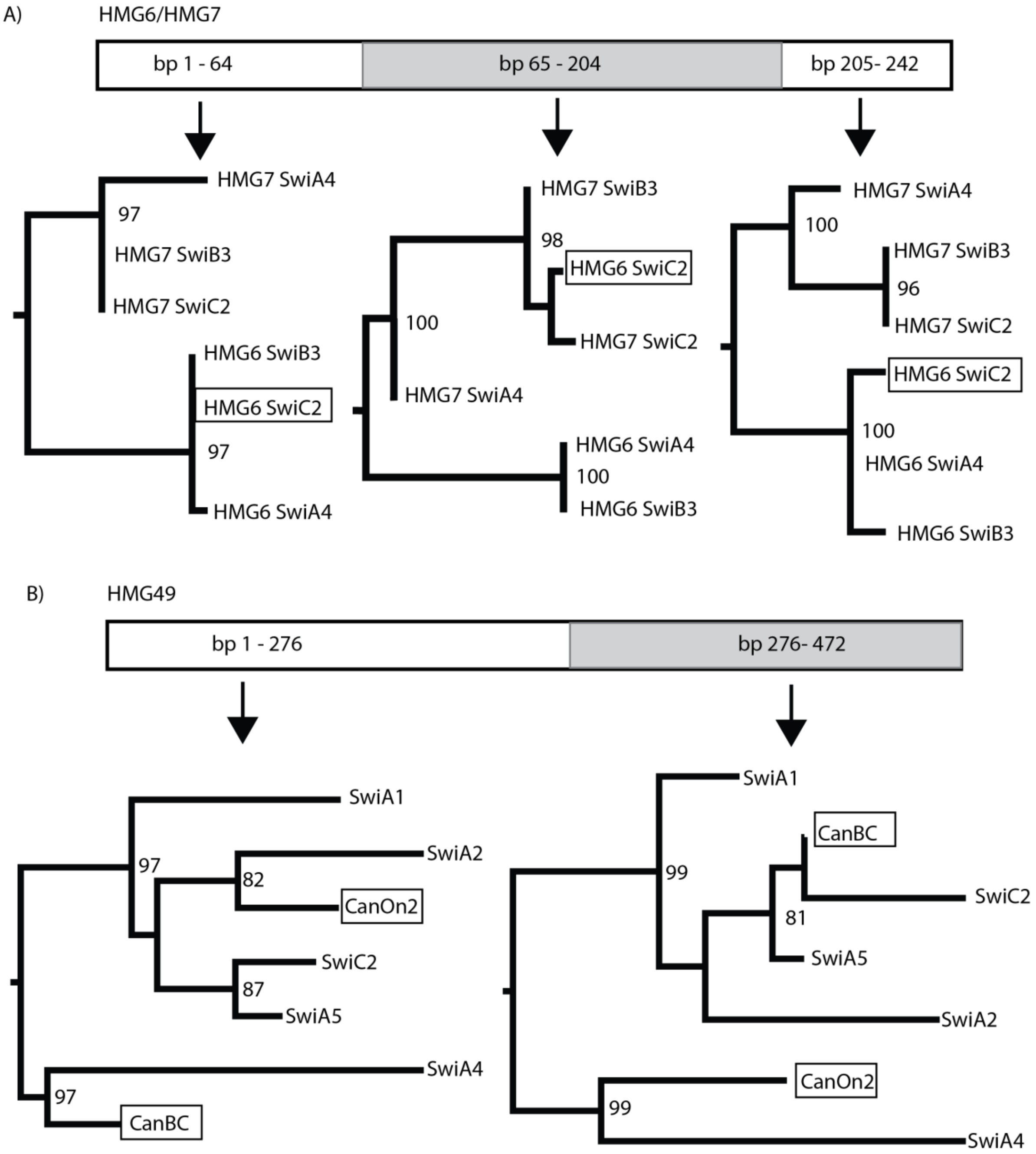


Figure 11. Phylogenetic incongruence of three sequence regions of HMG6 and HMG 7 from isolates SwiA4, SwiB3, and SwiC2 (a) and from two sequence regions of HMG49 from several isolates of AMF (b). Isolates outlined by boxes show significant (bootstrap ≥ 90) changes in position on the tree relative to other isolates. Bootstrap proportions over 90 are shown.

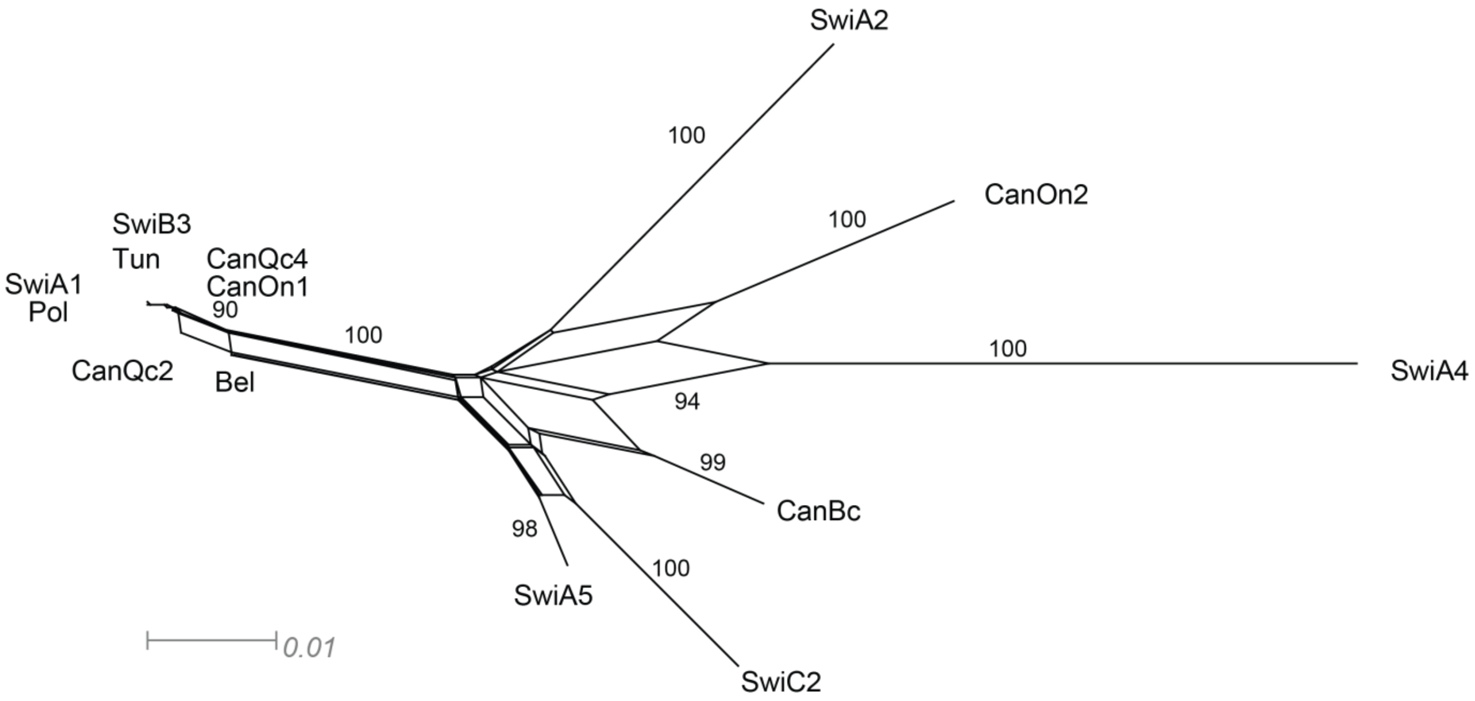


Figure 12. Neighbour-network of HMG49 in *R. irregularis* isolates of a broad geographic origine. Bootstrap proportion above 90% (n=1000 replicates) are shown. Isolate codes can be found in table 1. The scale bar represents 0.01 substitutions per site.

Genomic architecture surroundingsome MATA_HMG genes

The regions surrounding one the HMG6 were investigated using inverse PCR in order to identify a potential conservation in gene order (i.e. synteny) between this specific region of the AMF genome and that of known fungal mating type loci. The high AT content of the *Rhizophagus* DNA presented three challenges for inverse PCR. First, the steps involving primer annealing and DNA polymerase extension required temperatures lower than those generally used in standard PCRs. Second, fewer regions were available to design primers with even proportions of CG·AT , and, third, most restriction enzymes that recognize 6-bp sites include one CG pair. In AT-rich DNA, this significantly limits the chances of suitable cut sites that are separated by distances in the size range for amplification by PCR and, as result, restriction enzymes that recognize 4-bp sites were used in subsequent inverse PCRs to improve the chance of amplification.Because of the atypical features of *R. irregularis* genomic DNA, the inverse PCR procedures were successful to expand sequence information for only one region of the *R. irregularis* genome; a 7272 bp long DNA stretch that was to harbour an additional HMG domain located approximately 3kb upstream of HMG6. This newly identified HMG was identical to HMG5 which we had already recovered from the transcriptome.

The entire 7272bp region acquired through inverse-PCR and Sanger sequencing was then used as query against available genome data from *Rhizophagus irregularis* (kindly provided by the JGI) to further explore the architecture of this region of the AMF genome. This search resulted in the identification of an overlapping contig, extending the downstream portion of our previous contig by a total of 2370bp. Surprisingly, this additional sequence information was found to containan additional*MAT*-HMG gene, resulting in a single genomic region harbouring a totalof three tandem-repeated *MAT*-HMG genes.The assembly was confirmed using PCR with specific primers and Sanger sequencing.

In order to acquire first-hand knowledge about the evolution of a relatively large region of the AMF genome, we decided to amplify the portion of this genomic region containing the three tandem *MAT*-HMGs from other strains of this species; namely the isolates SwiA4, SwiB3 and SwiC2. Interestingly, the acquired sequences revealed the presence of extensive inter and intra-isolate polymorphism. Identical sites between isolates was found on average 82.5%, with sequence divergence mostly affecting non coding stretches of this genomic region. In some cases, PCR reactions resulted in the amplification of multiple versions of the same DNA region in the different isolates. Specific examples of this intra-individual variation are shown on Figure 13, and are highlighted by theisolates SwiC2 and SwiA4 having up to 4 copies of a particular region were found to be present in one isolate.

***MATA_HMG gene expression during hyphal interactions* in the AMF**

A total of 6 AMF *MAT*-HMG genes (HMG6, 1,37, 22,52, 65) were selected for analysis of gene expression using quantitative Real-Time PCR. Selection of the three variable gene targets in this study was based on successful optimization of primer sets. Invariable target HMG genes were randomly selected amongst the group of *MAT*-HMGs which were monomorphic amongst allisolates we analyzed. All attempts tomeasure variation in gene expression for the most divergent HMGs we identified among the core isolates produced inconsistent standard curves and were therefore discarded for downstream analyses. This is likely due to poor primer specificity on all isolates and unfortunately, alleles of these divergent HMGs were not variable enough to design allele specific primers.

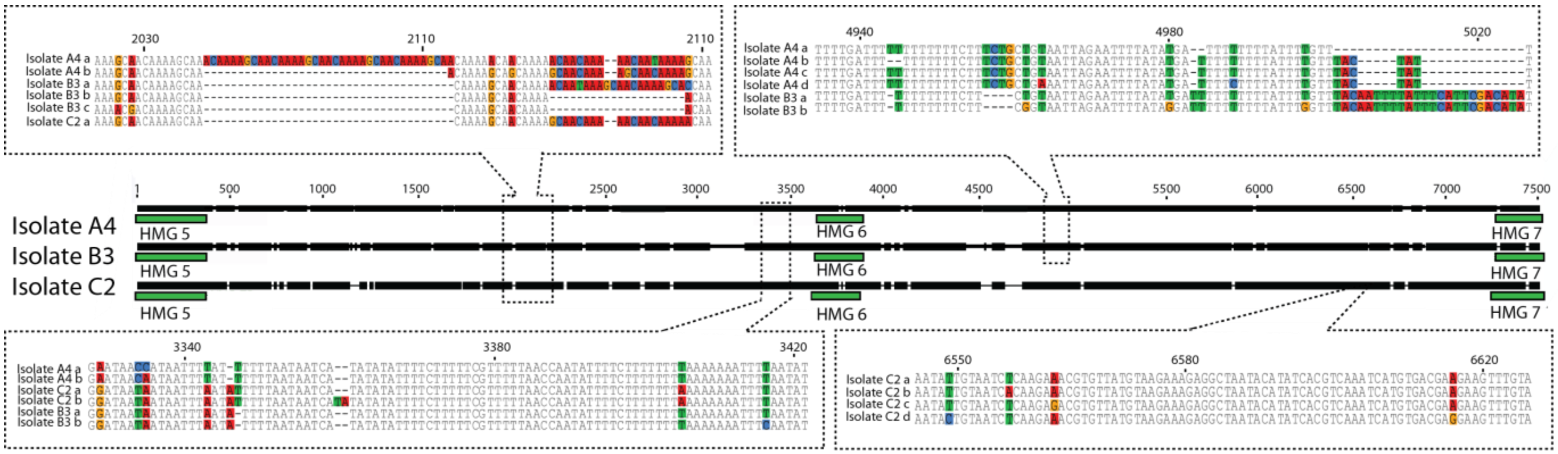


Figure 13. Tandem *MAT*-HMG locus amplified from isolates SwiA4, SwiB3 and SwiC2 of *R.irregularis*. The alignment of the tandem region in the three isolates is depicted in the middle of the figure. Green boxes represent the location of HMG5, HMG6, and HMG7 genes. Intraisolate variation at various sequence regions of the tandem MAT-HMG locus is shown within the hatched boxes and different versions of a sequence region recovered from an isolate is denoted by a different letter of the alphabet next to the isolates name. The scale above the alignments represents nucleotide location on the tandem MAT-HMG alignment.

A total of three HMGs harbouring significant divergence between the AMF isolates SwiA4, SwiB3 and SwiC2were selected for gene expression analyses following mycelial growth (HMG6, HMG1, HMG37). Importantly, HMG6 was found to be present as two highly divergent alleles (HMG6-A and B) and we were able to design allele-specific primers on each of them. These alleles were present in the isolate SwiA4 (HMG6-A) and SwiC2 (HMG6-B), respectively, while the isolate SwiB3 was found to harbour both alleles simultaneously (Figure 14). In parallel, a total of 3 MATA_HMG genes (HMG 22, HMG 52, HMG 65) showing no genetic divergence among isolates were also analysed. The expression of housekeeping genes encoding for the β-Tubulin and EF1α were used as a control and to normalize the RNA expression levels, respectively.

Among all the analysis we performed only a subset of genes were found to be significantly upregulated (Figure 15). Comparing isolates standalone conditions to all crosses involving the isolate, an increase in the expression level of all 4 variable gene targets was detected, of which 70% of these comparisons were found to be significant (Figure 16). Specifically, the expression levels of HMG6-A, HMG6-B, HMG 1 and HMG37 were found to increase on average by 12-fold following the interaction of hyphal networks. Interestingly, the expression of invariable gene targets (i.e. HMG 22, HMG 52, HMG 65, and β-Tubulin) did not change significantly when gene expression was compared between standalone to all-crosses.

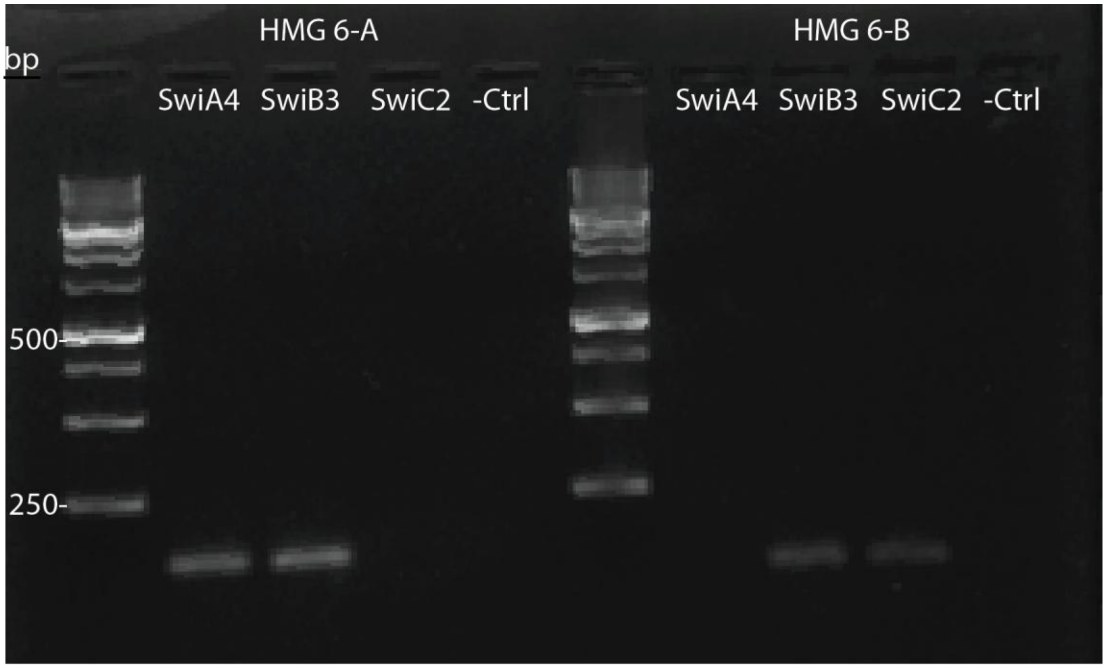
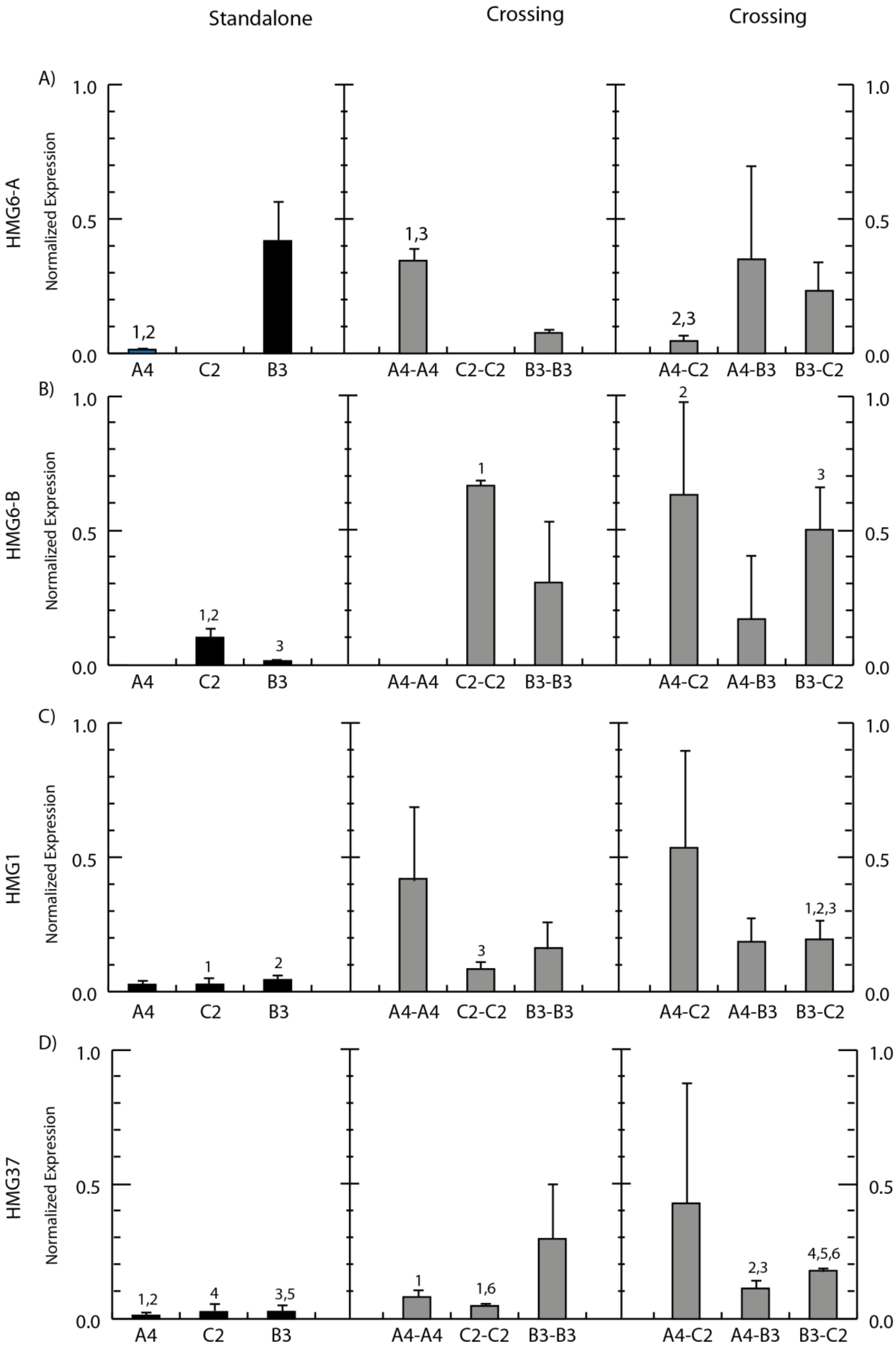


Figure 14. PCR products of HMG6 Q-RT-PCR amplicons for alleles HMG6-A and HMG6-B amplified from genomic DNA of isolates SwiA4, SwiB3 and SwiC2 of *R.irregularis* and a negative control (-Ctrl). Molecular sizes of thebands on the ladder are indicated on the left in base pairs (bp). Products were run on a 2% agros gel for 30 minutes at 90v.


Standalone
Crossing
Crossing
A)
HMG6-A
B)
HMG6-B
C)
HMG1
D)
HMG37
Normalized Expression
1.0
0.5
0.0
A4
C2
B3
A4-A4
C2-C2
B3-B3
A4-C2
A4-B3
B3-C2
1,2
1,3
2,3
1,2
3
1
2
3
1
2
3
1,2,3
1,2
4
3,5
1
1,6
2,3
4,5,6

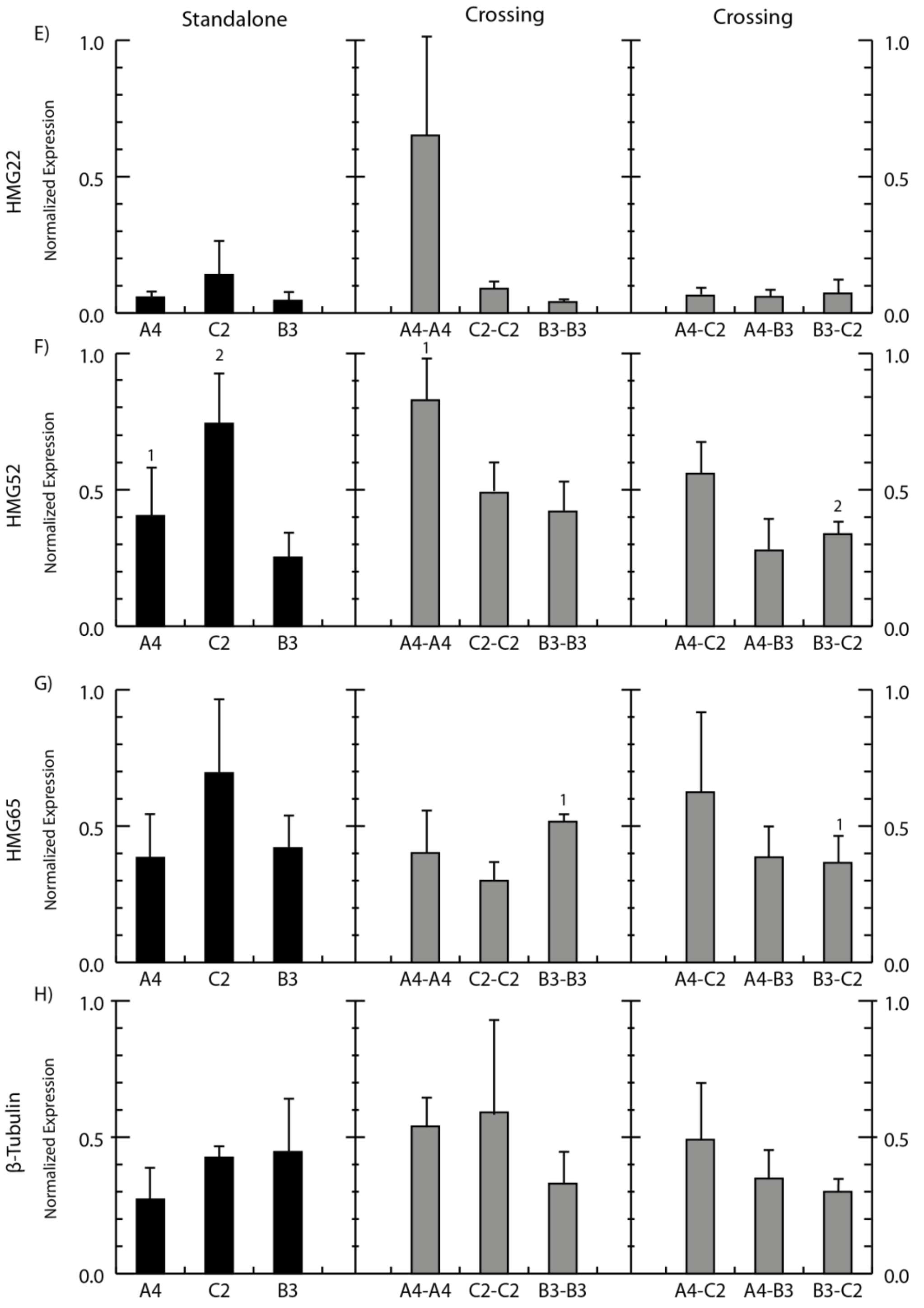


Figure 15. Gene expression of variable *MAT*-HMG domains (a) HMG 6-A, (b) HMG 6-B, (c) HMG 1, and (d) HMG 37 and invariable *MAT*-HMG targets (e) HMG 22, (f) HMG 52, (g) HMG 65, and a control target (h) β-tubulin in *R. irregularis* isolates SwiA4 (A4), SwiB3 (B3), and SwiC2 (C2) standalone conditions (black bars), and crossing conditions (grey bars): self- crossings A4-A4, C2-C2, and B3-B3, and out-crossings A4-C2 , A4-B3 and B3-C2. N=3 to 4 biological replicates were carried out for every experimental conidition. N=2 technical replicates were performed for all target gene/ experimental condition replicate except for reactions measuring the three variable gene target and β-tubulin in the standalone conditions (ie. standalone conditions A4, C2, and B3 measured with HMG49, 52, 65 and β-tubulin) where n=1 reaction for each biological replicate measured with these targets. Axes for each gene target are rescaled to the highest expressed sample (ie. the sample with the highest expression is a value of 1). Numbers above error bars represent pairs of P≤0.05 expression changes between standalone and crossing coniditions or selfcrosses and outcrosses.

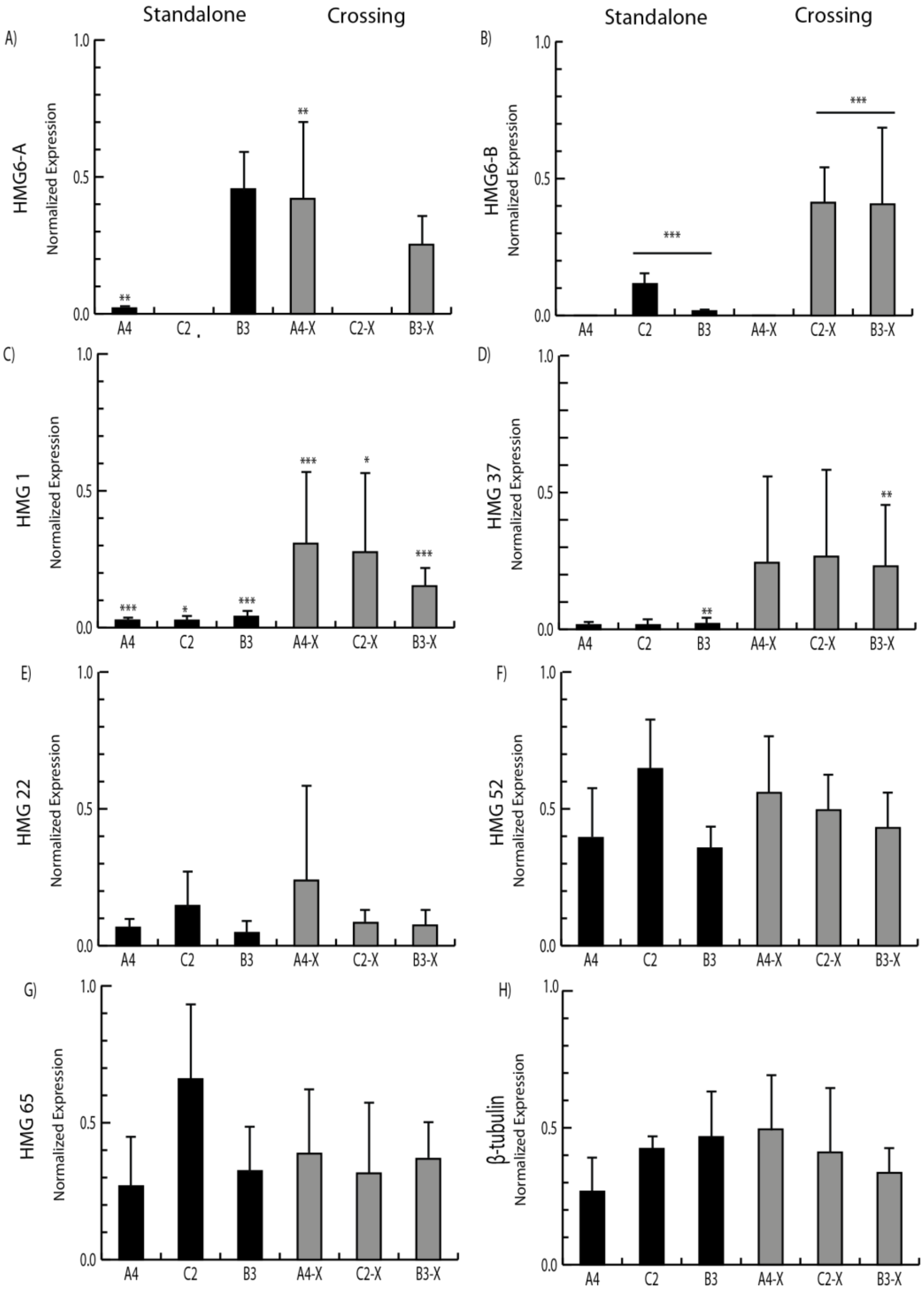


Figure 16. Gene expression of variable MAT-HMG domains (a) HMG6-A, (b) HMG6-B, (c) HMG1, and (d) HMG37 and invariable MAT-HMG targets (e) HMG22, (f) HMG52, (g) HMG65, and a control target (h) β-Tubulin in R. irregularis isolates SwiA4 (A4), SwiB3 (B3), and SwiC2 (C2) alone conditions (black bars) compared to crossings (grey bars) A4-X , C2-X , and B3-X, where X represents another isolate grown in the same crossing culture. Biological replicates were carried out for standalone conditions (n=3) and n=3 and crossing conditions in this comparison were pooled replicates of crosses involving a single isolate(ie. A4-X, C2-X, and B3-X; n=10 to 12 replicates each). N=2 technical replicates were performed for all reactions except for reactions measuring the three variable gene targets and β-Tubulin in the standalone conditions (ie. standalone conditions A4, C2, and B3 measured with HMG49, 52, 65 and β-tubulin) where n=1 reaction for each biological replicate measured with these targets. Axes for each gene target are rescaled to the highest expressed sample (ie. the sample with the highest expression is a value of 1). Asterisk's above bars represent levels of significance between crossings and standalone conditions: * $P \leq 0.05$, * $P \leq 0.02$, *** $P \leq 0.01$.

Discussion

Extensive gene duplications and the potential origins of intra-genomic variation in the AMF

In the present paper, searches for fungal *MAT*-HMGs across the transcriptome of two AMF species resulted in the identification of a surprisingly large number of gene homologues. Specifically, a total of 76 gene transcripts harbouring *MAT*-HMG genes were identified, all of which were later found to be present within the genomes of several isolates of *R. irregularis*. These *MAT*-HMG genes are all very diverse in sequence, as signified by the wide range of fungal *MAT*-HMGs with which these genes share homology, and appear to have rapidly expanded within the AMF lineage.

The extent of intra-genomic molecular diversity we report hereadds another level of complexity to what was previous known for other regions of the AMF genomes. In particular, a number of AMF genes have been known to be present in many “variants” indifferent individuals (i.e. Ribosomal RNA genes, POL1-like sequence (Clapp *et al.* 1999; Pawlowska and Taylor 2004)), but in no case was the amount of sequence variation found to be so inflated. Importantly, the identification of intra-individual molecular diversity in AMF has resulted in a long-standing debate of itsorigin. Specifically, a number of studies have proposed that genetic variation in AMF have arisen as a consequence of events of genome duplication (i.e. AMF are polyploid and homokaryotic), while other have suggested that the accumulation of mutations among genetically diverse and co-existing nuclei (i.e. heterokaryosis) has been the main source of genetic variability in this ecologically important group of fungi. (Hijri and Sanders 2005; Kuhn *et al.* 2001; Lanfranco *et al.* 1999;

LloydMacgilp *et al.* 1996; Pawlowska and Taylor 2004; Rodriguez *et al.* 2004; Stukenbrock and Rosendahl 2005).

Importantly, the present study helps to shed light onto the source of genetic diversity in the AMF, by providing a long-awaited view of the genome organisation along one 9.5kb regionof their genome. In particular, the identification of extensive tandem duplications of four *MAT*-HMGs in this study suggests that duplicative events may have played a major role in originating the extraordinaryintra-individual geneticvariation that is typical of these important plant symbionts. This would be especially true, if such molecular mechanisms were rampant across the genome of these organisms. The presence of inflated number of gene copiesin the AMF genome may also partly explain why attempts to assemble a meaningful sequence from representatives of this fungal group have failed to deliver at many levels.

Previous studies investigating recombination in AMF have reported its occurrence among isolates of single populations but in all cases these analysis have been based on analysis across multiple loci (Croll and Sanders 2009; den Bakker *et al.* 2010;Vandenkoornhuyse *et al.* 2001). Our study, however, indicates recombination has occurred within loci, occurring between individual loci located in a single genome and between a single locus located in different individuals.

Sequencing all *MAT*-HMGs identified in this study from representatives of many populations of *R. irregularis* have allowed us to explore the presence of recombination events in these supposedly ancient asexuals, and compare our data with that obtained along other regions of the AMF genomes, or other species. Interestingly, our results are in general agreement with those found by others (Croll and Sanders 2009; den Bakker *et al.* 2010;Vandenkoornhuyse *et al.* 2001), and also suggested the presence of recombination in this fungal group. Specifically, among a total of 76 *MAT*-HMGs, only three were found to harbour compelling signatures of recombination, supported by all our tests. Such signatures were notable, however, as they suggested the occurrence of homologous gene conversions not only between different paralogues of one strain, but also between strains of one species. Specifically, HMG6 from the isolate SwiC2 was found to contain an internal fragment that was highly similar to another gene located directly downstream on the same genome (HMG7), while sequences of the HMG 49 gene isolated from multiple *R. irregularis* isolates showed substantial evidence for recombination that could potentially follow genetic exchange between isolates of one species (i.e. the gene HMG49). So how did these events, which are typically linked with the presence of sexuality in other eukayotes, originate in the genomes of these supposedly ancient asexuals?

The origin of the recombining sequence we identified in the study (i.e. meiosis or mitosis) is difficult to assess and confounded by our lack of knowledge about these processes in AMF. Specifically, AMF have been proposed to be capable to undergo meiosis, but this has only been based on the presence of most meiosis-specific genes within their genomes, and a meiotic cyclehas never been formally observed in this fungal linage (Halary *et al.* 2011). Consequently, it is possible that meiosis-specific genes could potentially play a different functional role in the AMF that is not necessarily related with a sexual reproduction. Although the identification of potential gene shuffling between *R.irregularis* strains supports the idea that AMF are capable to exchange genetic material and undergo meiosis, we cannot exclude that such events could have arisen following non sexually-related processes (i.e. mitotic recombination; transposition).

Notwithstanding its origin, the detection of recombinant sequences in AMF in the present and past studies is essential for our understanding of the genetics of these organisms. In particular, it demonstrates that AMF have found a way to shuffle their genetic information within their genomes and, possibly, between different members of one species. It now remains to be seen whether AMF have been capable to do so by following a conventional sexual cycle (i.e. through meiotic recombination following nuclear fusion), or by using a number of non-sexual means that have yet to be formally described.

Are MAT-HMGs involved in sexually-related processes?

The identification of MAT-HMGs within different strains of the fungus *R. irregularis* is highly intriguing because these transcription factors are generally found within the mating type locus of other fungal organisms (Lee *et al.* 2010), and their identification raises the provocative suggestion that AMF may be using those genes for similar purposes (sex and partner recognition). Interestingly, those HMGs involved in partner recognition and mating all tend to follow common evolutionary patterns, including the presence of allelic variation between different members of one species that is sometimes conserved across species boundaries (Geiser 2008), and a relatively good conservation in the order of genes surrounding those*MAT*-HMG among members of one phylum (Lee *et al.* 2010). Moreover, *MAT*-HMGs that govern sexual identity and that are involved in sexual recognition are upregulated in some fungal lineages following crossing experiments involving members with different versions of these genes (Wetzel *et al.* 2012).

In the present study, searching for similar patterns of expression between AMF *MAT*-HMGs and homologues present at other fungal mating-types resulted in mixed results. In particular, gene expression of potential idiomorphs of *MAT*-HMGs did not always significantly change during all isolate crossing conditions. However, amongst all AMF *MAT*-HMGs there was a considerable difference in the expression of potential *MAT* idomorphs and invariable genes. Specifically, increases in gene expression were only observed in genes which exhibited allelic variation among isolates which also suggests that the best candidate mating-type genes for AMF are likely within the group of AMF *MAT*-HMGs which are divergent between isolates.

Two AMF *MAT*-HMG genes harboured two HMG domains on a single transcript which bears a resemblance to the structure of the *MAT*-HMG protein contained within the *MAT*-locus of several microsporidian species which also have two HMG domains present on a single gene (Lee *et al.* 2008). Some vertebrate HMGB proteins also contain two tandem HMG domains (Stott *et al.* 2006) but, like the microsporidian and AMF tandem HMG protein, the functional relevance of the tandem structure is unclear (Lee *et al.* 2008; Stott *et al.* 2006).

We searched the surrounding genomic region of one of the most divergent *MAT*-HMGs in our dataset to investigate gene order conservation between AMF and other fungal lineages. Although this investigation did not reveal a similar structure to the *MAT*locus of other fungi, we uncovered a genomic region in *R. irregularis* which consisted of a tandem array of three MAT-HMG genes. As far as we know, this genomic structure is entirely unique to the AMF lineage and possibly represents a small part of the highly duplicated nature of the AMF genome.

Concluding remarks

In the present study, the discovery of an expanded set of genes that are homologues to those found in fungal mating type loci provides long-awaited insights into the evolution of these ecologically relevant organisms. In particular, the identification in the AMF of additional genes known to be involved in sex in other fungi suggests that claims of long-term clonal evolution in these organisms may be more simplistic than previously thought and should be treated with caution. Similarly, the tremendously inflated copy number of these *MAT*-HMG genes in AMF provides additional evidence for the extreme complexity of their genomes. To this end, the discovery along the genome of *R. irregularis*of three MAT-HMGs with similar sequences repeated in tandem occurring between the genes highlights the potentially duplicated nature of this genome.

Overall, the present study representsa stepping stone for our understanding of the mode of propagation of these curious fungi,which certainly warrants further investigations at additional levels. For instance, because present searches were performed across transcriptome sequence data, it is likely that a large number of *MAT*-HMG genes have been overlooked because of their overall reduced, or condition-specific, expression levels. Ideally, the acquisition of large scale genome data will soon help to answer this specific question. Similar sequence data could also be used to acquire genome-wide expression level information using mRNA seq, a powerful approach that would allow the identification of the entire set of genes that are up-regulated following crossing experiments.

Conclusions

Arbuscular mycorrhizal fungi are of incredible importance to life on earth, but despite their tremendous ecological importance, we still lack a basic understanding of the extent of their reproductive modes. For many years, AMF have been considered as asexual organisms, but this has been only due to the lack of any observable sexual cycle and related morphological structures. The belief that AMF have evolved for extended periods of time, has resulted inthem being placed into a group of supposed eukaryotic aberrations, commonly referred to as ``ancient asexuals``. However, recent evidence for the presence of recombination, along with the presence of a core meiotic tool kit and two homologues of fungal *MAT*-HMG genes, has finally changed our perception of the productive mode that these organisms use to propagate. In particular, emerging evidence based on sequence data suggests that recombination may be more abundant in AMF than previously thought, possibly enabling the fungi to survive mutational loads for millions of years.

Consistent with this notion of sexuality in AMF, the data collected within this thesis demonstrates that AMF harbour a huge number of *MAT*-HMGgenes within their genomes, and that some of these could be implicated in processes associated with sexual reproduction.

Summary of findings

In this thesis, we identified ahighly expanded gene family of *MAT*-HMG genes in the genome of AMF, as we showed that alleles of these genes sometimes exhibit significant polymorphism among isolates of a population and, in some cases, this variation is conserved across species boundaries.This maintenance of allelic variation is generally linked with an evolutionary process called balancing selection, and which is known to have acted on *MAT* genes of other sexual fungi. Intriguingly, crossing isolates of *R. irregularis,* resulted in the upregulation of *MAT*-HMG gene expression - particularly for those showing extensive allelic variation. We also identified evidence of recombination among some AMF *MAT*-HMGs, and we propose that some of these events may have occurred between different isolates of one species; and that they therefore result from sexually related processes (i.e. nuclear exchange and fusion followed by mitosis or meiosis). Finally, our search for gene synteny between AMF *MAT*-HMGs and gene homologous of other fungi did not reveal gene order conservation; instead we found that three *MAT*-HMGs are located in tandem along a genomic region of approximately 9.5kb.

What might sex look like in AMF and future studies into sexual-related processes

AMF consist of massive ceonocytic networks where nuclei can flow freely through a common cytoplasm and in a homothallic scenario (i.e. having sex with oneself), nuclear fusion and meiosis could be occurring rampantly within a network of a single individual. Alternatively, considering the possibility that AMF are capable of heterothallism (i.e. having sex with different individual) a first step to sexual reproduction may be anastomosis between two different genotypes of an AMF species. Although anastomosis has been observed between genetically different individuals (Croll *et al.* 2009), the genes governing the process have not been discovered in AMF. In other fungi, chemical messengers (e.g. trisporic acid in the Zygomycota and a/α factor in some Ascomycota) between haphae facilitate recognition and fusion and similar molecules and pathways may play these roles in AMF (discussed in chapter 1). Presumably, anastomosis between genetically distinct individuals, would allow for nuclear exchange, although it has yet to be observed in AMF. Such an observation could be madewith through microscopic observation of nuclear movement during anastomosis possibly coupled with fluorescent labelling of different nucleotypes present in each anastamosing individual.

Nevertheless, karyogamy -or nuclear fusion- could be occurring, either within a cytoplasm of one individual in the case of homothallism, or following anastomosis and nuclear exchange in the case of heterothallism. Since no genes involved in nuclear fusion have been identified in AMF, searches across AMF sequence data for gene homologues involved in nuclear fusion in other fungi, such as the KAR genes in *S.cereveciae* (i.e. *KAR1, KAR2,KAR3,KAR4,KAR7,KAR8*) could be a first step in this investigation.

To complete a sexual cycle, meiosis must occur following karyogamy, but this also has yet to be observed. Investigation into if and when meiosis occurs, may be greatly facilitated by a complete genome sequence of an AMF, since it would enable the development of a greater number of genetic markers for studies aimed to detect recombination. In this case, a greater number of markers would result in a greater chance to detect recombination. Of special interest isif these searches would lead to the identification of a reciprocal crossover, a hallmark of conventional meiosis.

In the Kingdome fungi, the function of *MAT*-HMGs in the mating process can either be to control the production of pheromones which mediate sexual partner recognition, such as in some ascomycetes, or in the zygomycetes, may control the downstream mating processes, following pheromone production and communication (Wetzel *et al.* 2012). Whether the function of AMF *MAT*-HMGs resembles any one of these lineages is unknown but further study may reveal some clues as to their function.

Further investigation of the AMF MAT-HMG genes

Certainly, many more additional *MAT*-HMG genes exist within the genomes of the AMF isolates we surveyed, because our data is only based on those that are expressed in conventional in vitro cultures. Therefore, additional high throughput sequencing data, particularly genomic data, is necessary to reveal the true extent of this gene family in the AMF.

Presently, it is possible that one of or several of these *MAT*-HMGs are located within a regions corresponding to a b*ona-fide* AMF mating-type locus, but unfortunately none of the techniques we applied in this studies were sufficient to conclusively identify such a sexually-relevant region. In the mean time, a similar approach that we have taken in our study could be applied further to determine if additional AMF *MAT*-HMGs bear resemblances to *MAT*-loci of other fungi; individuals of AMF could be screened for polymorphic alleles, signatures of balancing selection could be searched, and genes could tested during crossing conditions for differential expression during interaction conditions involving two AMF hyphal networks. Given the large number of AMF *MAT*-HMGs, unravelling their function will be a tremendous task.

Unfortunately, in addition to problems arising from their atypical genome structure, functional studies of any genes, including MAT-HMGs, are very limited in AMF because of the lack of an appropriate transformation system. However, this problem could be partially circumvented by the use of Host Induced Gene Silencing (HIGS) (Nowara *et al.* 2010). In the case of AMF, this methodology has been successfully used to express RNAi constructs of a target transcript in the host carrot root, which can then silences the expression of that transcript in the AMF upon formation of the symbiosis (Helber *et al.* 2011). Additionally, heterelogous expression of AMF genes within fungal species, where transformation is much more efficient, could be used to investigate specific AMF gene function. For example, this approach has been used where an *STE12* gene mutant of *Colletotrichum lindemuthianum* was transformed with an AMF *STE12* homologue which successfully rescued the wild-type phenotype (Tollot *et al.* 2009).

These experiments may prove to be useful for the of the function of the different AMF *MAT*-HMG gene function. For example, does silencing of some *MAT*-HMGs effect the occurrence of anastomosis? On the other hand, genes that are associated with sexual processes such as those involved in meiosis and karyogamy may be investigated using such functional experiments as well. In the case of meiosis gene functional investigations, following the development of high-density markers along the genome to detect recombination, gene silencing of meiosis specific genes may be coupled with measurements of recombination frequency. In this case, changes in recombination frequency following gene silencing may provide clues as to meiosis gene function.

Implications

There has being a growing interest of the application of AMF to various fields, which all revolve around sustainable land use which includes ecologically sound agriculture and environmental remediation. This has been reflected by the growing number of companies commercially producing AMF inoculae for application in environmental practices. Since only specific isolates of AMF proliferate well under in-vitro conditions, the production of mass-inoculae has been limited to only a few selected isolates (mainly DAOM 197198); a feature that has severely confined the use of the AMF genetic diversity for environmental practices. For example, there is sufficient documentation to suggest different AMF genotypes are capable of exerting unique effects on their environments (i.e. differential effects on individual plant productivity, soil structure and nutrient use).

A complete understanding of the genetic basis of AMF partner recognition and compatibility (i.e. ability of two isolates to anastomosis) seems therefore necessary to for the development of an effective breeding program of these fungi, and this could be done using methodologies that are similar to those used for the conventional breeding of other organisms (i.e. plants, livestock and other fungi). The development of such a program would allow the introduction of novel traits into commercially viable inocula which may enable the propagation of a greater variety of AMF phenotypes for diverse applications.Within this context, the findings outlined in this thesis brings us one step closer to achieving this goal.

References

Alby, K., D. Schaefer and R. J. Bennett, 2009 Homothallic and heterothallic mating in the opportunistic pathogen Candida albicans. Nature 460**:** 890-893.

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

Bago, B., P. E. Pfeffer and Y. Shachar-Hill, 2000 Carbon metabolism and transport in arbuscular mycorrhizas. Plant Physiol 124**:** 949-958.

Banfield, J. F., W. W. Barker, S. A. Welch and A. Taunton, 1999 Biological impact on mineral dissolution: application of the lichen model to understanding mineral weathering in the rhizosphere. Proceedings of the National Academy of Sciences 96**:** 3404-3411.

Bennett, R. J., 2010 Coming of age--sexual reproduction in Candida species. PLoS Pathog 6**:** e1001155.

Bonfante, P., and A. Genre, 2010 Mechanisms underlying beneficial plant-fungus interactions in mycorrhizal symbiosis. Nature Communications 1**:** 48.

Bonfante, P., and M. A. Selosse, 2010 A glimpse into the past of land plants and of their mycorrhizal affairs: from fossils to evo-devo. New Phytologist 186**:** 267-270.

Boni, M. F., D. Posada and M. W. Feldman, 2007 An exact nonparametric method for inferring mosaic structure in sequence triplets. Genetics 176**:** 1035-1047.

Bruen, T. C., H. Philippe and D. Bryant, 2006 A Simple and Robust Statistical Test for Detecting the Presence of Recombination. Genetics 172**:** 2665-2681.

Burmester, A., M. Richter, K. Schultze, K. Voelz, D. Schachtschabel, W. Boland, J. Wostemeyer and C. Schimek, 2007 Cleavage of beta-carotene as the first step in sexual hormone synthesis in zygomycetes is mediated by a trisporic acid regulated beta-carotene oxygenase. Fungal Genet Biol 44**:** 1096-1108.

Bustin, S. A., 2005 Real-time polymerase chain reaction–towards a more reliable, accurate and relevant assay.

Butlin, R., I. Schon and K. Martens, 1998 Asexual reproduction in nonmarine ostracods. Heredity 81**:** 473-480.

Cavalier-Smith, T., 1998 A revised six-kingdom system of life. Biological Reviews 73**:** 203-266.

Clapp, J. P., A. H. Fitter and J. P. Young, 1999 Ribosomal small subunit sequence variation within spores of an arbuscular mycorrhizal fungus, Scutellospora sp. Mol Ecol 8**:** 915-921.

Corradi, N., and C. Charest, 2011 Some like it toxic. Mol Ecol 20**:** 3289-3290.

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**:** 366-369.

Corradi, N., and I. R. Sanders, 2006 Evolution of the P-type II ATPase gene family in the fungi and presence of structural genomic changes among isolates of Glomus intraradices. Bmc Evolutionary Biology 6.

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**:** 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.

Croll, D., L. Wille, H. A. Gamper, N. Mathimaran, P. J. Lammers, N. Corradi and I. R. Sanders, 2008 Genetic diversity and host plant preferences revealed by simple sequence repeat and mitochondrial markers in a population of the arbuscular mycorrhizal fungus Glomus intraradices. New Phytol 178**:** 672-687.

de la Providencia, I. E., F. A. de Souza, F. Fernandez, N. S. Delmas and S. Declerck, 2005 Arbuscular mycorrhizal fungi reveal distinct patterns of anastomosis formation and hyphal healing mechanisms between different phylogenic groups. New Phytologist 165**:** 261-271.

den Bakker, H. C., N. W. Vankuren, J. B. Morton and T. E. Pawlowska, 2010 Clonality and recombination in the life history of an asexual arbuscular mycorrhizal fungus. Mol Biol Evol 27**:** 2474-2486.

Deshmukh, S., R. Hückelhoven, P. Schäfer, J. Imani, M. Sharma, M. Weiss, F. Waller and K. H. Kogel, 2006 The root endophytic fungus Piriformospora indica requires host cell death for proliferation during mutualistic symbiosis with barley. Proceedings of the National Academy of Sciences 103**:** 18450-18457.

Dunthorn, M., and L. A. Katz, 2010 Secretive ciliates and putative asexuality in microbial eukaryotes. Trends in Microbiology 18**:** 183-188.

Dyer, P. S., and C. M. O'Gorman, 2011 A fungal sexual revolution: Aspergillus and Penicillium show the way. Current Opinion in Microbiology.

Edgar, R. C., 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32**:** 1792-1797.

Fang, W., and M. J. Bidochka, 2006 Expression of genes involved in germination, conidiogenesis and pathogenesis in Metarhizium anisopliae using quantitative real-time RT-PCR. Mycological Research 110**:** 1165-1171.

Francesconi, A., M. Kasai, R. Petraitiene, V. Petraitis, A. M. Kelaher, R. Schaufele, W. W. Hope, Y. R. Shea, J. Bacher and T. J. Walsh, 2006 Characterization and comparison of galactomannan enzyme immunoassay and quantitative real-time PCR assay for detection of Aspergillus fumigatus in bronchoalveolar lavage fluid from experimental invasive pulmonary aspergillosis. Journal of clinical microbiology 44**:** 2475-2480.

Friese, C. F., and M. F. Allen, 1991 The spread of VA mycorrhizal fungal hyphae in the soil: inoculum types and external hyphal architecture. Mycologia**:** 409-418.

Gehrig, H., A. Schussler and M. Kluge, 1996 Geosiphon pyriforme, a fungus forming endocytobiosis with Nostoc (cyanobacteria), is an ancestral member of the Glomales: evidence by SSU rRNA analysis. J Mol Evol 43**:** 71-81.

Geiser, D. M., 2008 Sexual structures in Aspergillus: morphology, importance and genomics. Medical mycology 47**:** 21-26.

Gibbs, M. J., J. S. Armstrong and A. J. Gibbs, 2000 Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics 16**:** 573-582.

Giovannetti, M., D. Azzolini and A. S. Citernesi, 1999 Anastomosis formation and nuclear and protoplasmic exchange in arbuscular mycorrhizal fungi. Appl Environ Microbiol 65**:** 5571-5575.

Gryganskyi, A. P., S. C. Lee, A. P. Litvintseva, M. E. Smith, G. Bonito, T. M. Porter, I. M. Anishchenko, J. Heitman and R. Vilgalys, 2010 Structure, function, and phylogeny of the mating locus in the Rhizopus oryzae complex. PLoS One 5**:** e15273.

Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk and O. Gascuel, 2010 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biology 59**:** 307-321.

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 Biol Evol 3**:** 950-958.

Harrier, L., 2001 The arbuscular mycorrhizal symbiosis: a molecular review of the fungal dimension. Journal of Experimental Botany 52**:** 469-478.

Harrison, M. J., 2005 Signaling in the arbuscular mycorrhizal symbiosis. Annu Rev Microbiol 59**:** 19-42.

Heitman, J., 2006 Sexual reproduction and the evolution of microbial pathogens. Current Biology 16**:** R711-725.

Heitman, J., 2010 Evolution of eukaryotic microbial pathogens via covert sexual reproduction. Cell Host Microbe 8**:** 86-99.

Helber, N., K. Wippel, N. Sauer, S. Schaarschmidt, B. Hause and N. Requena, 2011 A versatile monosaccharide transporter that operates in the arbuscular mycorrhizal fungus Glomus sp is crucial for the symbiotic relationship with plants. The Plant Cell Online 23**:** 3812-3823.

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

Hijri, M., and I. R. Sanders, 2004 The arbuscular mycorrhizal fungus Glomus intraradices is haploid and has a small genome size in the lower limit of eukaryotes. Fungal Genetics and Biology 41**:** 253-261.

Hijri, M., and I. R. Sanders, 2005 Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei. Nature 433**:** 160-163.

Hosny, M., V. Gianinazzi-Pearson and H. Dulieu, 1998 Nuclear DNA content of 11 fungal species in Glomales. Genome 41**:** 422-428.

Hull, C. M., and J. Heitman, 2002 Fungal mating: Candida albicans flips a switch to get in the mood. Current Biology 12**:** R782-784.

Hull, C. M., and A. D. Johnson, 1999 Identification of a mating type-like locus in the asexual pathogenic yeast Candida albicans. Science 285**:** 1271-1275.

Hull, C. M., R. M. Raisner and A. D. Johnson, 2000 Evidence for mating of the "asexual" yeast Candida albicans in a mammalian host. Science 289**:** 307-310.

Humphreys, C. P., P. J. Franks, M. Rees, M. I. Bidartondo, J. R. Leake and D. J. Beerling, 2010 Mutualistic mycorrhiza-like symbiosis in the most ancient group of land plants. Nature Communications 1.

Huson, D. H., 1998 SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14**:** 68-73.

Huson, D. H., and D. Bryant, 2006 Application of phylogenetic networks in evolutionary studies. Molecular biology and evolution 23**:** 254-267.

Idnurm, A., 2011 Sex determination in the first-described sexual fungus. Eukaryot Cell.

Idnurm, A., F. J. Walton, A. Floyd and J. Heitman, 2008 Identification of the sex genes in an early diverged fungus. Nature 451**:** 193-196.

Jansa, J., A. Mozafar and E. Frossard, 2005 Phosphorus Acquisition Strategies within Arbuscular Mycorrhizal Fungal Community of a Single Field Site. Plant and Soil 276**:** 163-176.

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**:** 1-16.

Johnson, N. C., J. H. Graham and F. A. Smith, 1997 Functioning of mycorrhizal associations along the mutualism-parasitism continuum. New Phytologist 135**:** 575-586.

Jones, S. K., Jr., and R. J. Bennett, 2011 Fungal mating pheromones: choreographing the dating game. Fungal Genet Biol 48**:** 668-676.

Judson, O. P., and B. B. Normark, 1996 Ancient asexual scandals. Trends Ecol Evol 11**:** 41-46.

Koch, A. M., D. Croll and I. R. Sanders, 2006 Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. Ecol Lett 9**:** 103-110.

Koch, A. M., G. Kuhn, P. Fontanillas, L. Fumagalli, J. Goudet and I. R. Sanders, 2004 High genetic variability and low local diversity in a population of arbuscular mycorrhizal fungi.

Proceedings of the National Academy of Sciences of the United States of America 101: 2369-2374.

Kottke, I., and M. Nebel, 2005 The evolution of mycorrhiza-like associations in liverworts: an update. New Phytologist 167: 330-334.

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

Landi, L., S. Murolo and G. Romanazzi, 2012 Colonization of Vitis spp. Wood by sGFP-Transformed Phaeomoniella chlamydospora, a Tracheomycotic Fungus Involved in Esca Disease. Phytopathology 102: 290-297.

Lanfranco, L., M. Delpero and P. Bonfante, 1999 Intrasporal variability of ribosomal sequences in the endomycorrhizal fungus Gigaspora margarita. Mol Ecol 8: 37-45.

Lee, S. C., N. Corradi, E. J. Byrnes, 3rd, S. Torres-Martinez, F. S. Dietrich, P. J. Keeling and J. Heitman, 2008 Microsporidia evolved from ancestral sexual fungi. Current Biology 18: 1675-1679.

Lee, S. C., M. Ni, W. J. Li, C. Shertz and J. Heitman, 2010 The Evolution of Sex: a Perspective from the Fungal Kingdom. Microbiology and Molecular Biology Reviews 74: 298-+.

Lee, S. C., L. M. Weiss and J. Heitman, 2009 Generation of genetic diversity in microsporidia via sexual reproduction and horizontal gene transfer. Commun Integr Biol 2: 414-417.

Li, C. H., M. Cervantes, D. J. Springer, T. Boekhout, R. M. Ruiz-Vazquez, S. R. Torres-Martinez, J. Heitman and S. C. Lee, 2011 Sporangiospore Size Dimorphism Is Linked to Virulence of <italic>Mucor circinelloides</italic>. PLoS Pathog 7: e1002086.

Lin, X., C. M. Hull and J. Heitman, 2005 Sexual reproduction between partners of the same mating type in Cryptococcus neoformans. Nature 434: 1017-1021.

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 Evolutionary Biology 9: 272.

Liu, Y. J., M. C. Hodson and B. D. Hall, 2006 Loss of the flagellum happened only once in the fungal lineage: phylogenetic structure of kingdom Fungi inferred from RNA polymerase II subunit genes. Bmc Evolutionary Biology 6: 74.

LloydMacgilp, S. A., S. M. Chambers, J. C. Dodd, A. H. Fitter, C. Walker and J. P. W. Young, 1996 Diversity of the ribosomal internal transcribed spacers within and among isolates of Glomus mosseae and related mycorrhizal fungi. New Phytologist 133: 103-111.

Logsdon Jr, J. M., 2008 Evolutionary Genetics: Sex Happens in Giardia. Current Biology 18: R66-R68.

Malik, S. B., A. W. Pightling, L. M. Stefaniak, A. M. Schurko and J. M. Logsdon, 2008 An Expanded Inventory of Conserved Meiotic Genes Provides Evidence for Sex in Trichomonas vaginalis. PLoS One 3.

Margulis, L., 1981 Endosymbiotic Theory. Symbiosis in Cell Evolution-life and its environment on the early earth: 1-14.

Martin, D., and E. Rybicki, 2000 RDP: detection of recombination amongst aligned sequences. Bioinformatics 16: 562-563.

Martin, D. P., P. Lemey, M. Lott, V. Moulton, D. Posada and P. Lefeuvre, 2010 RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26: 2462-2463.

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: 747-750.

Martin, F., A. Kohler, C. Murat, R. Balestrini, P. M. Coutinho, O. Jaillon, B. Montanini, E. Morin, B. Noel, R. Percudani, B. Porcel, A. Rubini, A. Amicucci, J. Amselem, V. Anthouard, S. Arcioni, F. Artiguenave, J. M. Aury, P. Ballario, A. Bolchi, A. Brenna, A. Brun, M. Buee, B. Cantarel, G. Chevalier, A. Couloux, C. Da Silva, F. Denoeud, S. Duplessis, S. Ghignone, B. Hilselberger, M. Iotti, B. Marcais, A. Mello, M. Miranda, G. Pacioni, H. Quesneville, C. Riccioni, R. Ruotolo, R. Splivallo, V. Stocchi, E. Tisserant, A. R. Viscomi, A. Zambonelli, E. Zampieri, B. Henrissat, M. H. Lebrun, F. Paolocci, P. Bonfante, S. Ottonello and P. Wincker, 2010 Perigord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. Nature 464**:** 1033-1038.

Martin, T., S. W. Lu, H. van Tilbeurgh, D. R. Ripoll, C. Dixelius, B. G. Turgeon and R. Debuchy, 2010 Tracing the origin of the fungal alpha1 domain places its ancestor in the HMG-box superfamily: implication for fungal mating-type evolution. PLoS One 5**:** e15199.

Metzenberg, R. L., and N. L. Glass, 1990 Mating type and mating strategies in Neurospora. Bioessays 12**:** 53-59.

Miller, R., J. Jastrow and D. Reinhardt, 1995 External hyphal production of vesicular-arbuscular mycorrhizal fungi in pasture and tallgrass prairie communities. Oecologia 103**:** 17-23.

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

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 29**:** 2199-2210.

Newsham, K. K., A. H. Fitter and A. R. Watkinson, 1995 Multi-functionality and biodiversity in arbuscular mycorrhizas. Trends Ecol Evol 10**:** 407-411.

Normark, B. B., O. P. Judson and N. A. Moran, 2003 Genomic signatures of ancient asexual lineages. Biological Journal of the Linnean Society 79**:** 69-84.

Nowara, D., A. Gay, C. Lacomme, J. Shaw, C. Ridout, D. Douchkov, G. Hensel, J. Kumlehn and P. Schweizer, 2010 HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen Blumeria graminis. The Plant Cell Online 22**:** 3130-3141.

O'Gorman, C. M., H. T. Fuller and P. S. Dyer, 2009 Discovery of a sexual cycle in the opportunistic fungal pathogen Aspergillus fumigatus. Nature 457**:** 471-474.

Padidam, M., S. Sawyer and C. M. Fauquet, 1999 Possible emergence of new geminiviruses by frequent recombination. Virology 265**:** 218-225.

Parfrey, L. W., D. J. G. Lahr and L. A. Katz, 2008 The dynamic nature of eukaryotic genomes. Molecular Biology and Evolution 25**:** 787-794.

Pawlowska, T. E., 2005 Genetic processes in arbuscular mycorrhizal fungi. FEMS Microbiol Lett 251**:** 185-192.

Pawlowska, T. E., and J. W. Taylor, 2004 Organization of genetic variation in individuals of arbuscular mycorrhizal fungi. Nature 427**:** 733-737.

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.

Perrin, N., 2012 WHAT USES ARE MATING TYPES? THE “DEVELOPMENTAL SWITCH” MODEL. Evolution 66**:** 947-956.

Poggeler, S., 2002 Genomic evidence for mating abilities in the asexual pathogen Aspergillus fumigates. Current Genetics 42**:** 153-160.

Posada, D., 2009 Selection of models of DNA evolution with jModelTest. Methods Mol Biol 537**:** 112.

Posada, D., and K. A. Crandall, 1998 Modeltest: testing the model of DNA substitution. Bioinformatics 14**:** 817-818.

Ramesh, M. A., S. B. Malik and J. M. Logsdon, 2005 A phylogenomic inventory of meiotic genes: Evidence for sex in Giardia and an early eukaryotic origin of meiosis. Current Biology 15**:** 185-191.

Ravnskov, S., and I. Jakobsen, 1995 Functional compatibility in arbuscular mycorrhizas measured as hyphal P transport to the plant. New Phytologist 129**:** 611-618.

Redecker, D., R. Kodner and L. E. Graham, 2000 Glomalean fungi from the Ordovician. Science 289**:** 1920-1921.

Reller, M. E., E. G. Clemens, S. E. Schachterle, G. A. Mtove, D. J. Sullivan and J. S. Dumler, 2011 Multiplex 5′ Nuclease-Quantitative PCR for Diagnosis of Relapsing Fever in a Large Tanzanian Cohort. Journal of clinical microbiology 49**:** 3245-3249.

Remy, W., T. N. Taylor, H. Hass and H. Kerp, 1994 Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proceedings of the National Academy of Sciences 91**:** 11841-11843.

Riley, R., and N. Corradi, 2013 Searching for clues of sexual reproduction in the genomes of arbuscular mycorrhizal fungi. Fungal Ecology 6**:** 44-49.

Rillig, M. C., 2004 Arbuscular mycorrhizae and terrestrial ecosystem processes. Ecology Letters 7**:** 740-754.

Rispail, N., and A. Di Pietro, 2010 The two-component histidine kinase Fhk1 controls stress adaptation and virulence of Fusarium oxysporum. Mol Plant Pathol 11**:** 395-407.

Rodriguez, A., J. P. Clapp and J. C. Dodd, 2004 Ribosomal RNA gene sequence diversity in arbuscular mycorrhizal fungi (Glomeromycota). Journal of Ecology 92**:** 986-989.

Rodriguez, R., and R. Redman, 2008 More than 400 million years of evolution and some plants still can't make it on their own: plant stress tolerance via fungal symbiosis. Journal of Experimental Botany 59**:** 1109-1114.

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

Rozen, S., and H. Skaletsky, 2000 Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132**:** 365-386.

Salminen, M. O., J. K. Carr, D. S. Burke and F. E. McCutchan, 1995 Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. AIDS research and human retroviruses 11**:** 1423-1425.

Sanders, I. R., 2002 Ecology and evolution of multigenomic arbuscular mycorrhizal fungi. Am Nat 160 Suppl 4**:** S128-141.

Sanders, I. R., 2003 Preference, specificity and cheating in the arbuscular mycorrhizal symbiosis. Trends Plant Sci 8**:** 143-145.

Sanders, I. R., and D. Croll, 2010 Arbuscular Mycorrhiza: The Challenge to Understand the Genetics of the Fungal Partner. Annual Review of Genetics, Vol 44 44**:** 271-292.

Sanders, I. R., and D. Croll, 2010 Arbuscular mycorrhiza: the challenge to understand the genetics of the fungal partner. Annu Rev Genet 44**:** 271-292.

Schmittgen, T. D., and K. J. Livak, 2008 Analyzing real-time PCR data by the comparative CT method. Nature protocols 3**:** 1101-1108.

Schneider, U. V., N. D. Mikkelsen, A. Lindqvist, L. M. Okkels, N. Jøhnk and G. Lisby, 2012 Improved Efficiency and Robustness in qPCR and Multiplex End-Point PCR by Twisted Intercalating Nucleic Acid Modified Primers. PLoS One 7**:** e38451.

Schurko, A. M., and J. M. Logsdon, 2008 Using a meiosis detection toolkit to investigate ancient asexual "scandals" and the evolution of sex. Bioessays 30**:** 579-589.

Schurko, A. M., M. Neiman and J. M. Logsdon, 2009 Signs of sex: what we know and how we know it. Trends in Ecology & Evolution 24**:** 208-217.

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

Sędzielewska, K. A., J. Fuchs, E. M. Temsch, K. Baronian, R. Watzke and G. Kunze, 2011 Estimation of the Glomus intraradices nuclear DNA content. New Phytologist 192**:** 794-797.

Sentandreu, V., N. Jiménez-Hernández, M. Torres-Puente, M. A. Bracho, A. Valero, M. J. Gosalbes, E. Ortega, A. Moya and F. González-Candelas, 2008 Evidence of recombination in intrapatient populations of hepatitis C virus. PLoS One 3**:** e3239.

Silberberg, G., K. Baruch and R. Navon, 2009 Detection of stable reference genes for real-time PCR analysis in schizophrenia and bipolar disorder. Anal Biochem 391**:** 91-97.

Smith, F. A., I. Jakobsen and S. E. Smith, 2000 Spatial differences in acquisition of soil phosphate between two arbuscular mycorrhizal fungi in symbiosis with Medicago truncatula. New Phytologist 147**:** 357-366.

Smith, J. M., 1986 Evolution - Contemplating Life without Sex. Nature 324**:** 300-301.

Smith, J. M., 1992 Analyzing the mosaic structure of genes. Journal of Molecular Evolution 34**:** 126-129.

Smith, S. E., D. J. Read and J. L. Harley, 1997 *Mycorrhizal symbiosis*. Academic Press, San Diego, Calif.

Stott, K., G. S. F. Tang, K. B. Lee and J. O. Thomas, 2006 Structure of a complex of tandem HMG boxes and DNA. Journal of molecular biology 360**:** 90-104.

Stukenbrock, E. H., and S. Rosendahl, 2005 Clonal diversity and population genetic structure of arbuscular mycorrhizal fungi (Glomus spp.) studied by multilocus genotyping of single spores. Mol Ecol 14**:** 743-752.

Taylor, S., M. Wakem, G. Dijkman, M. Alsarraj and M. Nguyen, 2010 A practical approach to RT-qPCR—publishing data that conform to the MIQE guidelines. Methods 50**:** S1-S5.

Tilman, D., D. Wedin and J. Knops, 1996 Productivity and sustainability influenced by biodiversity in grassland ecosystems. Nature 379**:** 718-720.

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. Azcon-Aguilar, G. Becard, P. Bonfante, M. J. Harrison, H. Kuster, P. Lammers, U. Paszkowski, N. Requena, S. A. Rensing, C. Roux, I. R. Sanders, Y. Shachar-Hill, G. Tuskan, J. P. 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 Phytol 193**:** 755-769.

Tollot, M., J. Wong Sak Hoi, D. Van Tuinen, C. Arnould, O. Chatagnier, B. Dumas, V. Gianinazzi-Pearson and P. M. A. Seddas, 2009 An STE12 gene identified in the mycorrhizal fungus Glomus intraradices restores infectivity of a hemibiotrophic plant pathogen. New Phytologist 181**:** 693-707.

van Aarle, I. M., G. Viennois, L. K. Amenc, M. V. Tatry, D. T. Luu and C. Plassard, 2007 Fluorescent in situ RT-PCR to visualise the expression of a phosphate transporter gene from an ectomycorrhizal fungus. Mycorrhiza 17**:** 487-494.

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

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

Vandenkoornhuyse, P., K. P. Ridgway, I. J. Watson, A. H. Fitter and J. P. Young, 2003 Co-existing grass species have distinctive arbuscular mycorrhizal communities. Mol Ecol 12**:** 3085-3095.

Wada, R., J. Maruyama, H. Yamaguchi, N. Yamamoto, Y. Wagu, M. Paoletti, D. B. Archer, P. S. Dyer and K. Kitamoto, 2012 Presence and functionality of mating type genes in the supposedly asexual filamentous fungus Aspergillus oryzae. Appl Environ Microbiol 78**:** 2819-2829.

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

Welch, D. M., and M. Meselson, 2000 Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. Science 288**:** 1211-1215.

Wetzel, J., A. Burmester, M. Kolbe and J. Wöstemeyer, 2012 The mating-related loci sexM and sexP of the zygomycetous fungus Mucor mucedo and their transcriptional regulation by trisporoid pheromones. Microbiology 158**:** 1016-1023.

Wong, S., M. A. Fares, W. Zimmermann, G. Butler and K. H. Wolfe, 2003 Evidence from comparative genomics for a complete sexual cycle in the 'asexual' pathogenic yeast Candida glabrata. Genome Biology 4.

Zhu, Y. G., and R. M. Miller, 2003 Carbon cycling by arbuscular mycorrhizal fungi in soil-plant systems. Trends Plant Sci 8**:** 407-409.