UNIL | Université de Lausanne

Faculté de Biologie et Médecine – Département d'écologie et d'évolution

Genetic variability in Arbuscular Mycorrhizal Fungi: Effect on gene transcription of *Oryza Sativa*

Thèse de doctorat ès science de la vie (PhD)

Présentée à la faculté de Biologie et de médecine de l'Université de Lausanne

par

Alexandre Colard

Diplômé en Biologie

Université d'Orsay, France

Jury

Prof. Alexandre Reymond, Président du jury

Prof. Ian R. Sanders, Directeur de Thèse

Prof. Didier Reinhardt, Expert externe

Prof. Marc Robinson Rechavi, Expert interne

Year : (2011)

Genetic variability in Arbuscular Mycorrhizal Fungi : Effect on gene trancription of *Oryza sativa*

Alexandre Colard

UNIL | Université de Lausanne

Faculté de biologie et de médecine

Ecole Doctorale

Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président	Monsieur	Prof. Alexandre **Reymond**
Directeur de thèse	Monsieur	Prof. Ian **Sanders**
Experts	Monsieur	Prof. Marc **Robinson-Rechavi**
	Monsieur	Prof. Didier **Reinhardt**

le Conseil de Faculté autorise l'impression de la thèse de

Monsieur Alexandre Colard

Master II GCDE Université de Paris XI

intitulée

Genetic variability in Arbuscular Mycorrhizal fungi: Effect on gene transcription of *Oryza Sativa*

Lausanne, le 4 novembre 2011

pour Le Doyen
de la Faculté de Biologie et de Médecine

Prof. Alexandre Reymond

Table of Contents

Remerciements

Je voudrais tout d'abord remercier Ian Sanders pour m'avoir donné l'opportunité et la chance de faire cette thèse dans son laboratoire. Il m'a mis dans les meilleurs conditions possible et m'a laisser la liberté de mener les expériences de mon choix en me poussant et me soutenant avec un enthousiasme certain. J'ai pu acquérir durant mon doctorat tant de connaissances qui me serviront dans le futur et que je n'aurais jamais pu acquérir si je n'avais pas fait cette thèse.

Je voudrais remercier mon Jury, messieurs Marc Robinson-Rechavi et Didier Reinhardt et Jan ver der Meer d'avoir accepté de juger mon travail. Leurs critiques et leurs conseils me seront très utile pour la suite.

Je n'oublie pas non plus tout le groupe Sanders, les anciens, Alexander, Beat, Caroline Daniel, Helene, Luca, Martine, Nicolas, Rui, Tania, ainsi que les nouveaux, Aurélien, Ivan, Jeremy, Jose, Lucas, Pawel et Tania. Merci a vous tous. Une mention spéciale à Auré et Caro pour leur aide de tous les instants dans le bureau et en dehors ainsi qu'à Tania pour son aide très précieuse en fin de thèse.

Je remercie également tout les gens qui m'ont aidé à mettre en place mes manips et ont répondu à toute mes questions même les moins pertinentes: le DAFL (et particulièrement Alexandra Paillusson et Manuel Bueno), Uta Paszkowski et Caroline Gutjahr, Francis Martin et son équipe à Nancy ainsi que la société académique vaudoise.

Et parce qu'une thèse ne se fait pas seul, je veux remercier tout les gens avec qui j'ai pu passer du temps, qui m'ont aidé, écouté et on fait la fête avec moi (et y'en a eu des belles !!). ils ont participé à mon bien être durant ces dernière années et m'ont permis de mener à bien cette thèse : Agnès, Alan, Aline, Annabelle, Aurélien, Biollay, Caro, Caroline, Céline, Céline, Charlène, Christine, Christine, David, Elodie, Farice, Fred, Godefroy, Guillaume, Helene, Joël, Laurie, Loïc, Lucie, Mika, Nadia, Nicolas, Pierre, Pierre, Roberto, Romain, Sonia, Tim, Thibault, Valou, Véro, Walid, mes potes du basket, du poker, tout le département d'écologie et d'évolution, tout les étudiants que j'ai eu l'occasion de croiser et tant d'autres personnes que j'oublie sans doute… Merci à vous tous.

Avec une mention spéciale aux trois personnes avec qui j'ai sans doute passé le plus de temps durant ces dernières années, Auré, Caro et Romain, avec qui j'ai fais les 400 coups…Merci les amis.

Je tiens également à remercier ceux qui était loin mais si proche en même temps et principalement Amira et Elodie qui m'ont énormément soutenu et aidé.

Je voudrais dire merci à ma Famille pour son aide depuis mon plus jeune âge mais surtout aussi pour avoir soutenu ma mère pendant ces dernières années quand elle en avait tant besoin.

Je remercie encore une fois ma meilleur amie, Caroline Angelard, sans qui cette thèse n'aurait sans doute jamais abouti tant elle m'a aidé scientifiquement mais aussi moralement. Merci a toi.

Je remercie pour finir ma maman qui a tout fait pendant toute sa vie pour que j'arrive aujourd'hui la ou j'en suis. Toujours à me soutenir et me pousser pour que j'aille plus loin même encore aujourd'hui. Je suis désolé de ne pas pu avoir été la tout les jours pour toi particulièrement lors de ces dernières années ou tu en avais le plus besoin et ou tu aurais du prendre plus soin de toi que de moi. Pour tout ça, je te remercie et je te dédie cette thèse.

Abstract

Arbuscular Mycorrhizal Fungi (AMF) form obligate symbioses with the majority of land plants. These fungi influence the diversity and productivity of plants. AMF are unusual organisms, harbouring genetically different nuclei in a common cytoplasm (known as heterokaryosis). Genetic variability has been shown between AMF individuals coming from the same population. Recent findings showed that genetic exchange between genetically different AMF individuals was possible. Additionnaly, segregation was shown to occur at spore formation in AMF. These two processes were shown to increase genetic variability between AMF individuals.

Because of the difficulty to study these organisms, almost nothing is known about the effect of intra-specific genetic variability in AMF on the plant transcriptome. The aim of this thesis was to bring insights into the effect of intra-specific genetic variability in AMF on plant gene transcription. We demonstrated that genetic exchange could influence expression of some symbiosis specific plant genes and the timing of the colonization of the fungi in plant roots. We also showed that segregation could have a large impact on plant gene transcription.

Taken together, these results demonstrated that AMF intra-specific variability could profoundly affect the life of plants by altering various molecular pathways. Moreover, results obtained on rice open a field of research on AMF genetics in impromvment of growth in agricultural plants and should be taken into account for future experiments.

Résumé

Les champignons endomycorhiziens arbusculaires (CEA) forment une symbiose obligatoire avec la majorité des plantes sur terre. Ces champignons peuvent influencer la diversité et la productivité des plantes avec lesquelles ils forment la symbiose. Les CEA sont des organismes particuliers de part le fait qu'ils possèdent des noyaux génétiquement différents (appelés hétérocaryosis) dans un cytoplasme commun. Il a été montré qu'il existait de la variabilité génétique intra-specific chez les CEA. De plus, des études recentes ont montré que l'échange génétique chez les CEA était possible entre des individus génétiquement différents tout comme la ségrégation qui a aussi été démontrée au moment de la formation des nouvelles spores chez les CEA. Ces deux processus ont été montrés comme pouvant créer aussi de la variabilité génétique intra-specific.

Du fait de la difficulté de travailler avec les CEA et à cause de la nouveauté de ces recherches, très peu de choses sont connues sur l'effet de l'échange génétique et de la ségrégation chez les CEA sur les plantes, et particulièrement au niveau moléculaire. Le but de cette thèse a été d'apporter la lumière sur les effets de la viariabilité génétique intra-specific chez les CEA, sur la transcription des gènes chez la plante. Nous avons pu montrer que l'échange génétique pouvait avoir des effets sur l'expression de gènes spécifiques à cette symbiose mais aussi pouvait influencer le timing de colonisation des racines de plantes par les CEA. Nous avons aussi montré que la ségrégation pouvait grandement influencer le transcriptome complet de la plante, et pas seulement les voies métaboliques spécifiques à la symbiose comme cela avait été montré auparavant.

L'ensemble de ces résultats démontre l'importance de la variation intra-specific chez les CEA sur les plantes et leur implication sur leur cycle de vie en changeant l'expression de voies métaboliques. De plus, ces résultats obtenus sur le riz ouvrent un champ de recherches sur les plantes destinées à l'agriculture et devraient être pris en compte pour des expériences futures.

Introduction

1. Symbiosis

Symbiosis is a close relationship between two organisms of different species. This link can be classified differently depending on the benefit of both partners. There are three principal types of symbiosis: Commensalism, parasitism and mutualism, but a continuum exists between these three types. First, commensalism happens when one organism has access to resources without giving anything back. Nevertheless, the host lives as normal, without any harm. Second, parasitism is observed when one organism obtains a gain whereas the partner, the host, is damaged. We find this type of symbiosis in almost all lineages, from bacteria to vertebrates. Third, mutualism is the case when both partners obtain benefits from the interaction. In particular, it can occur between plants and bacteria such as legumes and *rhizobium*. *Rhizobium* fixes nitrogen in nodules, and obtains carbon from the plant. This association can trigger the alteration of host gene expression (Yokota *et al.* 2011). Another example is the symbiosis between fungi and plants called mycorrhizal symbiosis. It is one of the most widespread symbioses on Earth. Three different types of mycorrhizal fungi exist: 1) Ecto-myccorrhizal fungi, where the fungi mostly belong to the *Basidiomycota*; 2) Ericoid mycorrhizal fungi, with mostly *Ascomycota* fungi; 3) Arbuscular mycorrhizal fungi (AMF). AMF represent the most abundant mycorrhizal symbiosis (Smith *et al.* 2008).

AMF are associated with more than 80% of land plants (Smith *et al.* 2008). Ribosomal deoxyribonucleic acid (rDNA) sequencing supported that AMF are in a monophyletic group: the *Glomeromycota* (Schussler *et al.* 2001). AMF-like fossil structures have been found in early Devonian plants such as *Aglaophyton major*. This species is known to have existed at least 400 millions years ago (Remy *et al.* 1994). Therefore, this symbiosis coincides with the apparition of the first terrestrial plants, suggesting that AMF have played a role in the colonization of land by plants (Selosse *et al.* 1998; Redecker *et al.* 2000; Brundrett 2002).

2. Life cycle of arbuscular mycorrhizal fungi

AMF form a network of mycelium formed by hyphae exclusively belowground. They produce spores to disperse. AMF are coenocytic, meaning that no septa (walls) are present in the hyphae. Therefore, nuclei are totally free in the spores and in the hyphae, and have the potential to move anywhere in the entire mycelial network. If we consider a multinucleate spore as an individual, we can describe their life cycle as follows (Figure 1): Spores are produced on the termini of hyphae in the soil. When they germinate, spores produce hyphae. The fungi then penetrate the root surface by forming an appressorium that will allow hyphae to enter into the cell. These intraradical hyphae will penetrate the cell using a structure called a "prepenetration apparatus" (Genre *et al.* 2005), which will guide the intraradical hyphae through epidermal cells toward the cortex. AMF, then, form arbuscules; Structures within cortical cells where the nutrients are exchanged between the plant host and the fungus. Some AMF species also form structures called vesicles inside the roots. Following colonization, the fungus produces new external hyphae in the soil which will take up nutrients and transport them back to the plant. Finally, AMF produce new spores that can germinate and colonize new hosts. A complete life cycle for AMF is at least 10-12 weeks for *Glomus intraradices* but it can take longer for other species. Since AMF are unspecific symbionts, they can form new symbioses with other plant species (Smith *et al.* 2008). Because they are coenocytic, resources and nuclei can potentially move in any direction within the mycelial network. Indeed, the movement of cytoplasm is bidirectional in hyphae as observed by Giovannetti *et al.* (1999) but it is not known how the flux is controlled. Movement of resources within the mycelial network and between the different hosts is relevant from an ecological point of view (Francis *et al.* 1984).

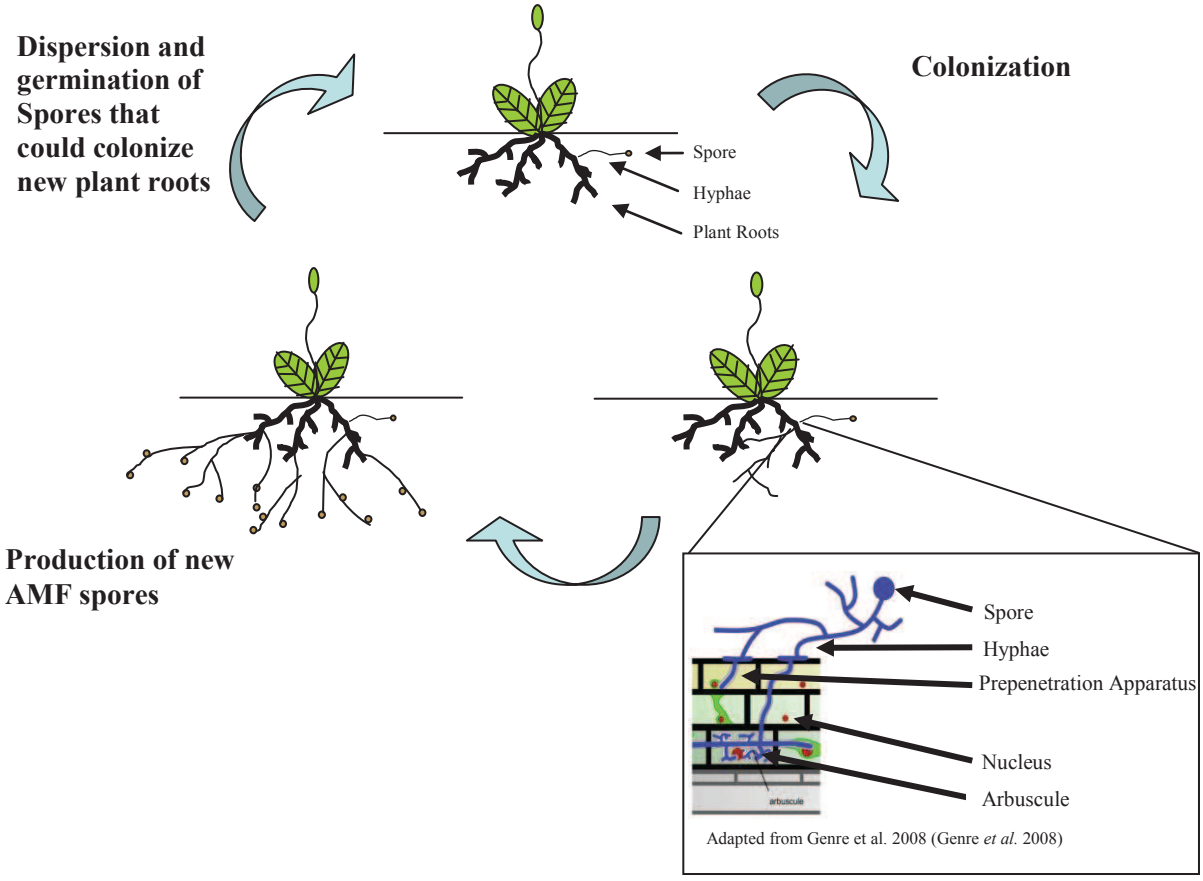


Figure 1: Life cycle of an AMF

3. AMF and ecology

AMF are obligate symbionts and cannot accomplish their life cycle without plants. This interaction is mediated through a structure, the arbuscule, inside the roots where nutrients are exchanged. AMF mainly supply plants with phosphate (Smith *et al.* 2008), but they also provide other resources such as nitrogen, zinc and copper (Harrison 1997; Harrison 1999; Hodge *et al.* 2001; Harrison *et al.* 2002; Smith *et al.* 2008). Moreover, AMF enhance plant resistance to different stresses, such as heavy metals (Ricken *et al.* 1996), herbivore pressure (Gange *et al.* 2003) and drought (Ebel *et al.* 1996).

In exchange, plants supply AMF with carbohydrates generated by photosynthesis (Jakobsen *et al.* 1990). It is also thought that different plants can exchange nutrients (eg. phosphate) via the underground fungal network (Francis *et al.* 1984; Frey *et al.* 1992; Newman *et al.* 1993).

Environmental conditions have been shown to have important consequences on the AM symbiosis. In some cases, the arbuscular mycorrhizal (AM) symbiosis may not be beneficial for plants, as cost on weights benefits, leading to growth depression (Peng *et al.* 1993; Al-Karaki 1998; Graham *et al.* 2000; Valentine *et al.* 2001). A study in 1997 postulated that AMF symbiosis cover a mutualism-parasitism continuum rather than being exclusively mutualistic (Johnson *et al.* 1997).

Since AMF are involved in a symbiosis with more than 80% of plants (Smith *et al.* 2008), they can influence plants species diversity (Grime *et al.* 1987) and can increase ecosystem productivity (van der Heijden *et al.* 1998; Klironomos *et al.* 2000). Thus, AMF are of central importance from an ecological and economic point of view (Rillig 2004).

Because they are ancient and associated with a large range of plant species, we can expect AMF to have evolved into many species. However, only 200 morphospecies have been described so far, based on phenotypic observations of spore morphology (Morton *et al.* 1990).

4. Genetics of AMF

Because of their ecological importance and their ubiquity on Earth, many studies have already been directed at understanding these organisms. However, it is fundamental to expand our knowledge at the molecular level, in order to understand how these organisms interact with plants. Since AMF are obligate symbiotic organisms, and therefore need plants to accomplish their life cycle, contaminations by plant products, or by micro-organisms living within the symbiont, have always been a critical problem for performing molecular experimental studies (Clapp *et al.* 1999). Moreover, the long duration of the fungal life cycle (10-12 weeks for *G. intraradices*, one of the fastest growing AMF species) and the low DNA quantity obtained after extraction are also major inconveniences. Successful cultivation of most AMF species is restricted to greenhouse cultures. However, even with the use of sterilized soil and plant seedlings, this kind of culture presents a great risk of contamination by non-AMF DNA. Nevertheless, *in vitro* systems allow us to obtain clean DNA by culturing the fungi with media on petri-dish with agrobacterium-transformed carrot roots (Becard *et al.* 1988). Basic features of genome architecture can be determined with the use of this system.

Most of the knowledge on AMF genomes comes from *G. intraradices* because this species can easily be cultivated *in vitro*. *G. intraradices* DNA harbours a low content of Guanine and Cystein (30 to 35% of the whole genome) (Hosny *et al.* 1997) against 50% for the majority of other organisms and particularly plants. This difference allows us to detect plant or micro-organism contaminations in the extracted DNA (Corradi *et al.* 2004). The size of the *G. intraradices* genome was estimated around 16.54 Mega base pairs (Mbp) by re-association kinetics (Hijri *et al.* 2004). The DNA content per nucleus was estimated at between 0.14 picogram (pg) for AMF *Scutellospora pellucida* to 1.15 pg for *Scutellospora gregaria* by flow cytometry (Hosny *et al.* 1998). The ploidy level is unknown for most AMF species, although *G. intraradices* and *Scutellospora castanea* appear to be haploid, based on reassociation kinetics (Hijri *et al.* 2004). AMF spores can also contain different numbers of nuclei depending on the AMF species, from a minimum of 1000 nuclei per spore for *Glomus caledonium* (Burggraaf *et al.* 1989) up to 35000 nuclei for *Gigaspora decipiens* (Viera *et al.* 1990). But these numbers should be taken with precaution because of the cultivation difficulties and the report of possible contaminations by other micro-organisms (Bianciotto *et al.* 1996; Hijri *et al.* 2002).

Observation of many different ribosomal gene sequences within AMF spores has lead to the hypothesis that genetically different nuclei could coexist within AMF individual indicating that AMF could be heterokaryotic (Sanders *et al.* 1995; Sanders *et al.* 1996). In 2001, Kuhn *et al.* (2001) showed that nuclei are different within spores of *Scutellospora castanea*, showing an intra-individual genetic variability in AMF and supporting the hypothesis that AMF are heterokaryotic. The AMF heterokaryosis hypothesis has generated strong debates during the last decade. A study proposed that high polymorphism in AMF was due to high ploidy rather than heterokaryosis (Pawlowska *et al.* 2004). They showed that *Glomus etunicatum* contained 13 variants of a polymerase-like gene (POL). After single spore culture, they extracted DNA from five spores and carried out polymerisation chain reaction (PCR) on each. They were able to amplify the 13 POL variants in each extraction supporting their hypothesis. Bever and Wang (2005) responded by showing hyphal fusion between different individual could maintain the high polymorphism in AMF in a heterokaryotic system. Moreover, Hijri *et al.* (2005) demonstrated that the ploidy of *G. etunicatum* should be at 13N if the high ploidy was the reason of the high polymorphism. DNA content of *G. etunicatum* nucleus was measured and was found to be 37.45 Mbp per nucleus. Therefore, the genome size of this fungus would

be 2.88Mb if the ploidy level was 13N, which is highly unlikely for a eukaryote genome. The heterokaryotic hypothesis was also supported by another study that showed the haploid status of two other AMF species, *G. intraradices* and *S. castanea* (Hijri *et al.* 2004). This study refutes the hypothesis that genetic variation within this fungus is due to very high ploidy. Today, it is generally accepted that AMF are heterokaryotic because of the unusually high polymorphism seen in the genome sequence (Martin *et al.* 2008). Dikaryotes, composed of alls containing two different nuclei, has already been known for a long time in many fungi, but this state is only transitory (Okazaki *et al.* 2008). Other cases of heterokaryosis are known in the phylum *Ascomycota* (McGuire *et al.* 2005). Heterokaryosis in AMF implies many consequences and raises new questions about sexuality in AMF and the transfer of their genetic information to the next generation.

5. Asexuality

5.1 Asexuality

By definition, "sex" in eukaryotes is accomplished by meiosis, which is one of the most conserved processes in evolution (Kleckner 1996). Meiosis begins with one diploid cell and produces four haploid cells called gametes. Evolutionary theory predicted that asexual species would accumulate deleterious mutations in their genome and then should become extinct (Smith 1998).

Nevertheless, AMF are considered to have been asexual for 400 million years. Neither meiosis nor any sexual structures have been observed in AMF, except in a controversial study which described sexual structures like zygospores, but this observation has never been reproduced (Tommerup *et al.* 1990).

Some studies suggested that recombination occurred in highly polymorphic sequences of gene encoding for a binding protein (BiP) and rDNA genes of *G. intraradices* (Gandolfi *et al.* 2003). Futhermore, a recent study has shown that recombination has occurred in a Swiss AMF population of *G. intraradices* using a multi-locus sequence analysis (Croll *et al.* 2009). Nevertheless, direct evidence for meiosis has never been shown in AMF.

There are many fungi from different clades that were considered asexual until genetic tools revealed that the meiosis machinery was present and functional. For example, *Aspergillus*

fumigatus (a human pathogen from *ascomycota* phylum (Dyer *et al.* 2005) was described as asexual, until its genome was fully sequenced in 2005 (Gow 2005). Then, it was shown in 2009 that indeed it could undergo meiosis *in vitro* and produced cleistothecia (that are spherical fruiting bodies characteristic of sex in the aspergilli) with a specific medium at 30°C in the dark (O'Gorman *et al.* 2009).

5.2. Genetic exchange (Figure 2)

The fusion of hypha, called anastomosis, was shown between hyphae of the same individual in AMF (Giovannetti *et al.* 1999). Then in 2009 it was shown that anastomosis between individual hyphae of genetically different isolates was possible (Croll *et al.* 2009). The authors identified specific single sequence repeat (SSR) markers for each of two parental lines and revealed the presence of both parental specific markers in the same spore obtained after the fusion. Furthermore, the offspring exhibit new phenotypes and produce new viable spores that are able to colonize new plants. They suggested that this mix of nuclei between isolates could be considered as parasexuality, even if no recombination between different nuclei occurs.

The exchange of nuclei following the fusion of hyphae of two different AMF individuals provides an explanation for the high intra-individual genetic diversity found in *G. intraradices*.

5.3. Segregation (Figure 2)

Segregation was shown in 2010 in the AMF *G. intraradice*s (Angelard *et al.* 2010). In this study, they showed that SSR markers were in different proportions in new spores than in the mother spores. It was also shown that spores with different combinations of nuclei had varying degrees of success in their root colonization ability and differentially affect plant growth and the transcription of symbiosis specific genes rice (Angelard *et al.* 2010; Angelard *et al.* 2011).

In conclusion, these two processes (genetic exchange and segregation) produce new progeny with novel genetic content and symbiotic effects. However, these processes are still largely unknown, as well as their consequences on the host.

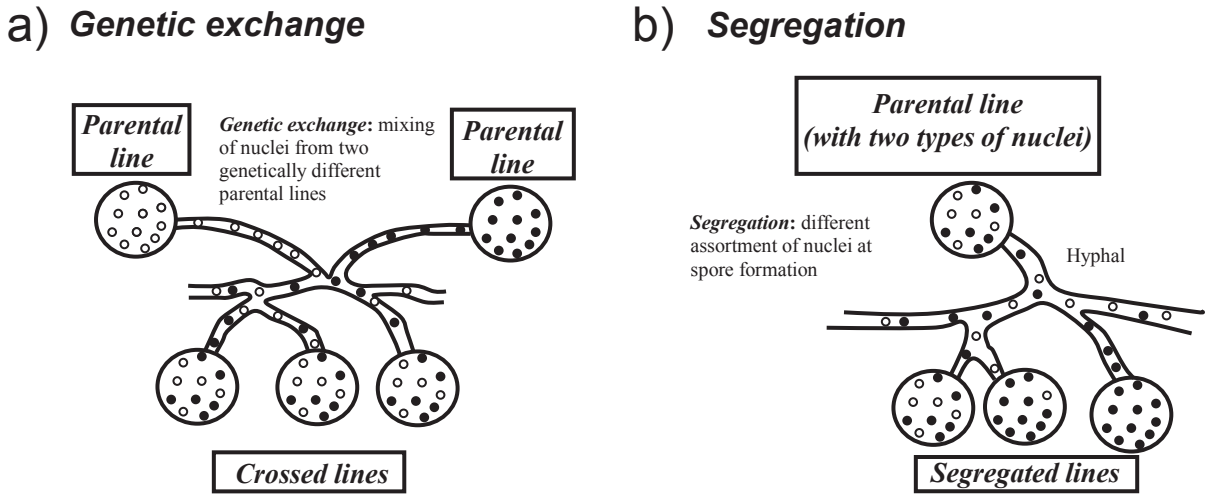


Figure 2: Schematic representation of two processes genetic exchange and segregation.

a) Genetic exchange: Hyphae of two genetically different spores can fuse, allowing the mixing of the nuclei.

b) Segregation: Genetically different nuclei can be differentially segregated in spores.

Illustration from Angelard *et al. (2010)*

6. Genomic features

In 2004, the first sequencing of a whole AMF genome was initiated by the American Department of Energy's Joint Genome Institute (JGI). They chose the AMF *G. intraradices* for several reasons:

1) This species is one of the most ubiquitous AMF species present in both temperate and tropical ecosystems.

2) It is important economically since it colonizes many plant species, including agriculturally important species (wheat, alfalfa or rice) and some model plants used in research (*Medicago truncatula* or *Lotus japonicum*).

3) As mentioned above, *G. intraradices* colonizes plants quickly in comparison with some other AMF species and it is simple to culture it *in vitro*. For industries, it is the prime species cultured for commercial purposes. It is the only AMF species for which spores are available in "pure" form and in large quantities for commercial purposes.

About 80 Mbp of the *G. intraradices* genome has been sequenced using Sanger methods or the 454 GS-2 pyrosequencing technique. Some laboratories have also contributed by adding

sequences (Martin *et al.* 2008). The 80 Mbp available should cover the whole genome since it represents five times the estimated size. But the whole genome sequence is still not assembled because of two problems: First, heterokaryosis leads to different haplotypes and could be associated with different alleles and second, the presence of numerous short repeated sequences increases the risk of misassemblies.
A recent assembly of 52.5 Mbp found a small genome size of 16 Mbp. But the genome is probably larger than the amount of DNA per nucleus because of the different haplotypes present in one individual, and now it is predicted to be around 150 Mbp.

The sequencing of the mitochondrial genome was easier. The mitochondrial DNA was less polymorphic and moreover a standard set of mitochondrial genes were found. Collaboration between the University of Montreal and the University of York published the entire mitochondrial genome of *G. intraradices* in 2009 (Lang *et al.* 2009).

Today, large banks of expressed sequence tags (EST) and cDNA libraries exist for *G. intraradices*. These sequences come from germinated spores and extra-radical mycelium (Martin *et al.* 2008). A total of 83 539 putative transcripts were obtained. This number is high, again because of the high polymorphism present in the AMF. Only 7 750 genes where identified by sequence homology, while the other transcripts remained unknown. Soon, the transcriptome of *G. intraradices* will be released (E Tisserant *et al.*, *in press*)

7. AMF and Rice

Rice (*Oryza sativa*) is one of the most cultivated cereals on Earth. Like 80% of species plant, it can form symbioses with AMF. Rice constitutes about 20% of the daily caloric intake of the human population, and represents more than 50% of the calories consumed in Asian countries (Caicedo *et al.* 2007). In the past 30 years, world rice production has doubled, but it is still insufficient to cope with the increasing global demands (Fischer *et al.* 2000; Sasaki *et al.* 2000). The world demand in rice is expected to increase at a rate of 1% per year from now until 2025. Beyond its importance as the worlds first crop, rice is also an excellent model for plant genomics (Izawa *et al.* 1996). Among cereals, rice has the smallest genome with an estimated size of 430 Mbp as compared to the significantly large genome sizes of sorghum,

maize, barley, or wheat (about 750, 3000, 5000, and 16000 Mbp, respectively). According to different studies, the number of genes of *O. sativa* is between 30 000 and 50 000 (Temnykh *et al.* 2000; Goff *et al.* 2002). Many of these genes are automatically annotated and almost all of them are mapped on microarrays.

The relation between rice and AMF is not so well documented. Given the importance of rice there are markably few published studies on rice and AMF. Some arbuscular mycorrhizal symbiosis specific genes and symbiosis pathways have started to be studied in rice (Paszkowski *et al.* 2002; Guimil *et al.* 2005; Gutjahr *et al.* 2008; Gutjahr *et al.* 2009). These studies showed the influence of AMF on its host and its effect at the molecular level. For example, Guimil *et al.* (2005) compared the transcriptome profile of rice inoculated with *G. intraradices* and uninoculated rice. They showed that 224 genes were specifically affected by the presence of AMF. However, the effects of the genetic variability between individuals produced by genetic exchange and segregation in AMF on rice are poorly known and the investigation of these two processes can bring a lot of information about the interaction between AMF and rice.

8. Aims and Outlines of the thesis

In this introduction, I have listed several features AMF in general and of *G. intraradices* in particular.

The discovery of the genetic exchange and segregation processes in AMF opens the possibility of a new research field because it allows the researcher to alter the genetic make-up of the fungus and see its effects on the symbiosis and how it affects plant gene expression.

In 2010, Angelard *et al.* (2010) investigated the effect of these two mechanisms on the two plant species *Plantago lanceolata* and *O. sativa*. They found that such processes had a strong effect on plant traits, such as dry mass, but also influences the fungal traits such as colonization rate. Nevertheless, no test was done to check the effect of such a mechanism at the molecular level.

In chapter 1 of this thesis, I investigated the effects of the genetic exchange in *G. intraradices* on some *O. sativa* symbiosis specific genes. The aim of the study was to show that genetic exchange in AMF could have important consequences for their symbiotic effects on plants. Using a set of arbuscular mycorrhizal specific genes previously characterized (Gutjahr *et al.*

2008), I carried out quantitative PCR on retro-transcribed rice root RNA that were inoculated with parental lines or crossed-lines. Additionally, I investigated whether the colonization dynamics could be affected by genetic exchange.

In 2004, it has been shown that AMF lines coming from the same population have a high genetic variability (Koch *et al.* 2004). Moreover, in 2006, it has been demonstrated that this genetic variability within species can cause variation in plant growth (Koch *et al.* 2006). In 2005, a study investigated the effects of two different AMF species, *G. intraradices* and *Glomus mosseae*, on the transcriptome of *Medicago truncatula* (Hohnjec *et al.* 2005). They found that the variability among AMF species could alter plant transcriptome profile. However, the effects of this intra-specific genetic variability on the plant transcriptome remain unknown. In chapter 2, I investigated the consequences of genetic variability between two *G. intraradices* isolates lines coming from the same population on the whole rice gene transcriptome using microarrays. Furthermore, this approach allowed me to identify genes mycorrhizal that are sensitive to the mycorrhizal symbiosis that were not detected in other studies.

Angelard *et al.* (2010) found that segregation occurred in AMF and could have an effect on rice growth and on symbiosis specific gene transcription (annexe I). Some segregated lines increased rice growth by a factor of five, while other lines did not influence plant growth. The effects of segregation in AMF on the rest of the rice transcriptome remained unknown. In Chapter 3, I investigated, for the first time, the effect of segregation in *G. intraradices* on the whole rice transcriptome using segregated lines previously characterized by Angelard *et al.* (2010). My study has shown the potential of segregation in AMF on the plant host transcriptome.

The annexe chapters present further research conduct during my thesis. The first provides evidence for copy number polymorphism within a population of *G. intraradices* (Corradi *et al.* 2007). The second presents the molecular evolution of a potentially important gene in symbiosis (Corradi *et al.* 2009). The third provides molecular evidence for the presence and transcription of meiosis specific genes in *G. intraradices*. The last annexe presents a study showing that segregation can occur in AMF. This study also investigates the effect of segregation on plant growth and on rice symbiosis specific genes (Angelard *et al.* 2010).

Authors contributions

Chapter 1: C.A. performed the statistical analysis and helped to draft the manuscript, A.C. inoculated the plants, measured plant growth and colonization, performed the RNA extraction and the quantitative PCR. A.C. and C.A. designed the experiment. A.C. and I.S. wrote the manuscript. All authors discussed the results and commented the manuscript.

Chapter 2: C.A. inoculated the plants, performed the statistical analysis and helped to draft the manuscript. J.M. and A.C. did the root RNA extraction, carried out the microarray analysis, performed and analysed the quantitative PCR. A.C. and I.S. wrote the manuscript. All authors discussed the results and commented the manuscript.

Chapter 3: C.A. inoculated the plants, performed the statistical analysis and helped to draft the manuscript. J.M. and A.C did the root RNA extraction, performed and analysed the quantitative PCR. A.C. carried out the microarray analysis. A.C. and I.S. wrote the manuscript. All authors discussed the results and commented the manuscript.

Chapter 1

Genetic exchange in an arbuscular mycorrhizal fungus increases rice growth and alters mycorrhizal specific gene transcription

Alexandre Colard, Caroline Angelard and Ian R. Sanders

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

This chapter is *in press* in *Applied and Environmental Microbiology*

Abstract

Arbuscular mycorrhizal fungi (AMF) are obligate symbionts with most terrestrial plants. They improve plant nutrition, particularly phosphate acquisition and are, thus, able to improve plant growth. In exchange, the fungi obtain photosynthetically fixed carbon. AMF are coenocytic, meaning that many nuclei co-exist in a common cytoplasm. Genetic exchange has recently been demonstrated in AMF, allowing nuclei of different AMF to mix. Such genetic exchange was previously shown to have negative effects on plant growth and alter fungal colonization. However, no attempt was made to detect whether genetic exchange in AMF can alter plant gene expression, and if this effect was time dependant. Here, we show that genetic exchange in AMF can also be beneficial for rice growth and that symbiosis specific gene transcription is also altered by genetic exchange. Moreover, our results show that genetic exchange can also change the dynamics of the colonization of the fungus in the plant. Our results demonstrate that simple manipulation of the genetics of AMF can have important consequences for their symbiotic effects on plants such as rice; the most globally important crop. Exploiting natural AMF genetic variation, by generating novel AMF genotypes through genetic exchange, is potentially a useful tool in the development of AMF inocula that are more beneficial for crop growth.

Introduction

Plants and arbuscular mycorrhizal fungi (AMF) probably form the most common symbiosis on earth. The fungi have repeatedly been shown to confer a number of benefits to plants including increased growth through improved nutrient acquisition, especially phosphate (Smith *et al.* 2008). In return, the fungal partner receives photosynthetically fixed carbon (Jakobsen *et al.* 1990). Mycorrhizal fungal diversity is also a factor determining plant diversity and productivity (van der Heijden *et al.* 1998; van der Heijden *et al.* 1998; Klironomos *et al.* 2000). AMF were thought to be asexual and indeed they can reproduce clonally (Judson *et al.* 1996). They are also coenocytic, and thus, nuclei co-exist in a common cytoplasm. Some AMF have been shown to be heterokaryotic, containing populations of genetically different nuclei within one fungus (Kuhn *et al.* 2001; Hijri *et al.* 2005). The consequences of this heterokaryosis are still poorly understood for either the fungi or the plants with which they form symbioses.

Although it is well understood that AMF diversity is important in ecosystems, little is known about the role of AMF genetic variation in shaping plant growth. Recently, two processes have been shown in the AMF *Glomus intraradices* that generate genetically novel AMF; namely genetic exchange and segregation, that give rise to AMF lines (named crossed lines or segregated lines, respectively) that are genetically and phenotypically distinct from their parents (named parental lines) (Croll *et al.* 2009; Angelard *et al.* 2010). Following segregation, segregated lines were shown to have different effects on rice growth and also induced changes in transcript levels of rice genes that are known to be specifically transcribed during the mycorrhizal symbiosis (Angelard *et al.* 2010). Some crossed lines suppressed the growth of rice compared to inoculation with parental lines. More recently, both crossed and segregated AMF lines were also shown to exhibit different levels of colonization in rice roots, although this was only measured at one time point, 12 weeks after inoculation (Angelard *et al.* 2011).

Knowledge about the molecular pathway of arbuscular mycorrhiza (AM) establishment is mostly derived from studies in legumes, but recently, AM-specific gene transcription in rice was described by Gutjahr *et al.* (2008). They showed which genes are specifically transcribed during the establishment and development of the symbiosis and, for some genes, also where they are transcribed. They characterized 18 genes that are specifically transcribed in the

mycorrhizal symbiosis, and tested their transcription at 3, 5, 7 and 9 weeks after inoculation. The function of most of these genes was unknown. Four of them (AM1, AM2, AM3 and AM11) were transcribed very early during the colonization, 3 weeks after inoculation. The other AM-specific genes started to be transcribed at 7 weeks after inoculation (Table 1). The transcription of some of these genes was subsequently shown to be altered by segregation in AMF at one time point, 12 weeks after inoculation (Angelard *et al.* 2010). Indeed, segregated AMF lines induced different transcription levels of some of these genes compared to rice inoculated with the crossed lines from which they originated.

Given that we know genetic exchange in AMF can alter colonization levels in rice and that genetic changes in AMF, through segregation, can also have an effect on transcription of symbiosis specific-genes, it is pertinent to investigate whether genetic exchange in AMF affects transcription of symbiosis specific genes. In particular, we asked whether any observed transcriptional changes in rice are affected by genetic exchange in AMF during the first weeks of establishment of the symbiosis, whether transcriptional changes are related to colonization levels and plant growth and whether they are the same genes as those altered by segregation in AMF. Based on previous studies (Angelard *et al.* 2010; Angelard *et al.* 2011), we hypothesized that rice exhibits different levels of symbiosis specific gene transcription by being inoculated with crossed AMF lines compared to being inoculated with parental AMF lines, and that those differences would change through time.

We chose two genetically different parental lines of *G. intraradices* and three AMF lines arising from crosses between the two parental lines to determine, the effect of genetic exchange in AMF on mycorrhizal specific gene transcription in rice. We also measured plant growth and fungal colonization to detect any correlations between phenotype and AM-specific transcription levels. Furthermore, we performed the experiment as a time course making measurements at 6, 9 and 12 weeks after inoculation with AMF. The time course gave us information about the dynamics of colonization among AMF lines and also the mycorrhizal specific plant gene transcription over time.

Table 1: Genes characterized by Gutjahr *et al.* (2008) and their temporal transcription.

Genes	Transcription			Putative function
	Early (< 3 weeks)	Late (> 7weeks)	Systemic	
AM1*	X			Putative classe III peroxidase (Prx53)
AM3*	X		X	Contains peptidoglycan binding LysM domain 1
AM10*		X		Similarity to putative hypersensitivity-related (Hsr) protein
AM11*	X			Hypothetical protein, similarity to nucleoid DNA-binding protein cnd41
AM14*		X		Serine-threonine kinase like
AM15*		X		Contains peptidoglycan bonding LysM domain 2
AM20*		X		Similar to AB-hydrolase associated lipase region
AM24*		X		Putative cDNA
AM25*		X		Putative MIP aquaporin, nodulin 26-like
AM26*		X		Serine-threonine kinase, calcium dependent (EF hand)
AM29*		X		Similar to Ring-H2 zinc finger protein-like
PT11*		X		PT11 phosphate transporter

Material and methods

Origin of the fungal lines

Glomus intraradices lines C3 and D1 were used in this study. They originate from one agricultural field in Tänikon, Switzerland (Koch *et al.* 2004). These lines were isolated from two different plots each separated by 60m. These two lines were previously shown to be genetically distinct, to differ in their phenotypic traits and to differentially affect plant growth (Koch *et al.* 2004; Koch *et al.* 2006). In a previous study, isolates C3 and D1 were co-cultured together to obtain crossed lines Sb, Sc2 and Sd which exhibited bi-parental inheritance (Croll *et al.* 2009). Sc2 in the present study corresponds to Sc in the study of Croll *et al.* (2009). Parental and crossed lines were cultivated in standard conditions with Ri T-DNA transformed carrot roots prior to inoculation on plants (Becard *et al.* 1988).

Plant growth conditions, inoculation with AMF and harvesting

The substrate for growing plants was composed of 60% sand and 40% loam and was autoclaved twice at 120° Celsius (C) and left open to the air for two weeks in order to oxidize any toxic compounds. The seeds of rice cv. Nipponbare (kindly supplied by Uta Patzowski's group, University of Lausanne) were surface-sterilized in Sodium hypochlorite NaClO 2% for 10 minutes, washed repeatedly in sterile water and placed on filter paper for germination. Eight days after germination, seedlings were put on trays with moist vermiculite, and were watered for another five days. AMF spores were extracted from four month old cultures using the same method as in Koch *et al.* (Koch *et al.* 2004). Each pot was filled with 150 millilitres (ml) of soil and one rice seedling was planted in each pot. Each plant was inoculated with 500 spores of one of the AMF crossed or parental lines. Some plants were not inoculated and served as the non-mycorrhizal control. Five AMF lines were used, plus an uninoculated plant treatment with 12 replicates for each of four sequential harvests giving a total of 288 pots. Plants were grown in a phytochamber with 60% humidity, 12 hours (h) light and 26°C during day and 22°C at night. Plants were watered 3 times a week with 10 ml of tap water. After 3 weeks, 10 ml fertilizer without phosphorus (mixture of 1 gramme (g) sequestrene (Syngenta) and 0.5 Hauert-Flovy Type A (Hauert) in 1 litre (l) of water) was given twice a week. Twelve plants inoculated with each AMF line and the uninoculated plants were harvested from the

phytochamber every 3 weeks for a period of 12 weeks, resulting in four harvest times, at 3, 6, 9 and 12 weeks after inoculation. The shoots of each plant were harvested, dried at 80°C for two days and weighed. The roots were immediately washed and separated into two parts. One third was stored in 50% ethanol in order to measure fungal colonization. The other two thirds were frozen in liquid nitrogen for subsequent RNA extraction and stored at -80°C. No plant died during the experiment. No AMF colonization could be detected in 88 plants that had been inoculated with AMF. These eighty-eight plants were randomly distributed among all treatments. This left 200 individuals that were used for further analyses.

Measurement of fungal colonization of roots

Root samples of each plant were stained with trypan blue as described by Munkvold *et al.* (2004). Fungal colonization (i.e. vesicular, arbuscular and hyphal colonization) was measured following the method of Mcgonigle *et al.* (1990). Fungal colonization, as defined in Mcgonigle *et al.* (1990), corresponds to the percentage of arbuscules plus the percentage of vesicles plus the percentage of hyphae only (without vesicles or arbuscules). Because the amount of roots was too low at 3 weeks after inoculation, we only checked for the presence or absence of the fungi inside roots but the percentage of root length colonized by AMF was not measured. All the non-mycorrhizal (NM) plants were checked and found to be AMF free.

RNA extraction and quantitative PCR

RNA extractions were performed directly on frozen roots. Roots of plants harvested at 3 and 6 weeks after inoculation were pooled (among samples of the same AMF treatment and from the same time point) because not enough material was available per plant. Roots were ground with a pestle in liquid nitrogen. We used the RNeasy Plant mini kit (Qiagen Inc®) to extract RNA following the manufacturer's recommendations. Precipitation with LiCl was also performed (Ambion®). DNase treatment was performed with DNase I grad (Invitrogen®). A first PCR on RNA was performed to amplify the housekeeping genes (cyclophilin and ubiquitin) to be sure that there was no DNA contamination, with the following conditions: 94°C for 5 minutes, 35 cycles [94°C 20 seconds, 58°C 20 seconds, 72°C 50 seconds] 72°C for 7 minutes. A reverse-transcription with Superscript III enzyme (Invitrogen®) with Poly dT primer was then performed following the instructions of the manufacturer. Each sample was

diluted 5 times. Quantitative PCRs were performed on 12 genes (see Table 2 for primer sequences) that are known to be up-regulated during the first weeks of colonization. Quantitative PCR conditions and the mix for each tube was identical to that described by Gutjahr *et al.* (2008) These genes had been found with Affymetrix microarray screening (Guimil *et al.* 2005) and previously checked by quantitative PCR (Gutjahr *et al.* 2008) except for OsPt11 gene, which was already characterized (Paszkowski *et al.* 2002). Power Sybr® green was used as the label (Applied Biosystems) for quantitative PCR and each sample was run in triplicate. Quantitative PCR was performed on Prism AB7900 thermocycler. Cyclophilin and ubiquitin transcript levels were used to normalize the data. Transcription values were calculated with SDS2.2.1 software (AB) and qBASE 1.3.5 software (Hellemans *et al.* 2007).

Statistical analyses

Plant growth

To meet the requirements of the statistical tests, Box-Cox transformations were used when needed so that data displayed normal distributions. To test whether the overall growth of NM plants differed from AMF inoculated plants, we pooled the data obtained for each harvest and each AMF treatment and we compared this with the NM treatment using one-way ANOVA. We found a strong effect of AMF inoculation on plant growth ($F_{1,173} = 263.27$, $p = 0.016$) with mycorrhizal plants growing larger than non-mycorrhizal plants. The NM treatment was then removed because it could bias the results of the statistical analyses used to identify whether growth differences existed among plants inoculated with different AMF. We performed a crossed two-way ANOVA with “harvest time” and “AMF line” as fixed factors.

Fungal colonization

The effect of the harvest time and of the AMF treatment on fungal colonization was also tested in a crossed two-way ANOVA with “harvest time” and “AMF line” as fixed factors. One-way ANOVAs were then performed for each harvest time separately. When ANOVAs were significant, we carried out a multiple comparison test using Tukey-Kramer HSD (Honestly significant difference) test with an alpha level of 0.05.

Plant gene transcription

Analyses on transcription levels were performed for each gene separately. The effects of the harvest time and of the AMF treatment on rice gene transcription were tested in crossed two-way ANOVAs with "harvest" and "AMF treatment" as fixed factors. One-way ANOVAs were then performed for each harvest separately, and when ANOVAs were significant, we carried-out a multiple comparison test using Tukey-kramer HSD test with an alpha level of 0.05.

All analyses were performed with the statistical program JMP® version 5.0 (SAS Institute Inc., Cary, NC, USA).

Table 2: Primer sequences used in this study and designed by Gutjahr *et al.*(Gutjahr *et al.* 2008)

Gene	Primer Sequences	5' → 3'
*Os*PT11	Forward:	GAGAAGTTCCCTGCTTCAAGCA
	Reverse:	CATATCCCAGATGAGCGTATCATG
*Os*AM1	Forward:	ACCTCGCCAAAATATATGTATGCTATT
	Reverse:	TTTGCTTGCCACACGTTTTAA
*Os*AM3	Forward:	CTGTTGTTACATCTACGAATAAGGAGAAG
	Reverse:	CAACTCTGGCCGGCAAGT
*Os*AM10	Forward:	AGAACACTTGTGGCCGTACTATAAGA
	Reverse:	CCTCTCGACGAAAGTACGGACTA
*Os*AM11	Forward:	TGAACGAAGACAGCAATACATCAA
	Reverse:	CGATCGATGGATTCATACTTCAGT
*Os*AM14	Forward:	CCAACACCGTTGCAAGTACAATAC
	Reverse:	GCACTTTGAAATTGGACTGTAAGAAA
*Os*AM15	Forward:	TCCGGCGCCACATAGTG
	Reverse:	TCCGTCGCACACGAGAAG
*Os*AM20	Forward:	TTTGGAAGGAACACTACTGGAGAAT
	Reverse:	CCGAAATCTAGTTTCGACAATGATT
*Os*AM24	Forward:	TCTTCATCACCGCCGACAT
	Reverse:	CGGCGAGATAGTGAGCATAAAGA
*Os*AM25	Forward:	CTTGCTGCCTTCCTCTATGGA
	Reverse:	CGAGAAGTCGACGACTCCTACAC
*Os*AM26	Forward:	GGTTGTTGCGGCATGTGTAC
	Reverse:	AGCCATGTCCCTAGCGAGGTA
*Os*AM29	Forward:	TGCGACGTGATCAGCCAC
	Reverse:	TGCACGCACCTGTTCCAC
*Os*Cyclophilin	Forward:	GTGGTGTTAGTCTTTTTATGAGTTCGT
	Reverse:	ACCAAACCATGGGCGATCT
*Os*Ubiquitin	Forward:	CATGGAGCTGCTGCTGTTCTAG
	Reverse:	CAGACAACCATAGCTCCATTGG

Results

Shoot dry weight & fungal colonization

In this experiment, inoculated plants (averaged over all mycorrhizal treatments) grew larger than uninoculated plants. Plant shoot dry weight also significantly differed among rice inoculated with different AMF lines, with the lowest values mostly obtained for plants inoculated with the two parental lines and the highest values obtained for plants inoculated with the crossed lines Sc2 or Sd (Figure 1; Table 3). The significant difference among shoot dry weights of plants inoculated with parental and crossed lines occurred throughout the experiment (Figure 1; Table 3). There was no significant interaction between AMF line and harvest time on shoot dry weight (Figure 1; Table 3).

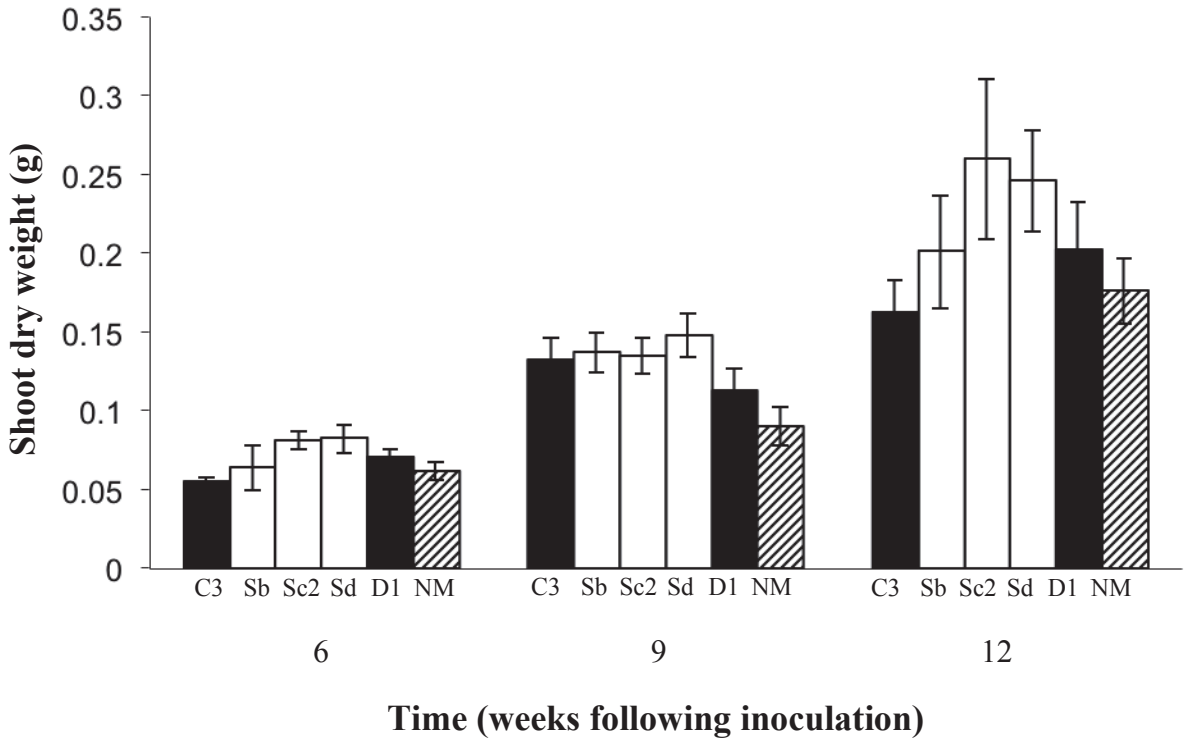


Figure 1: Plant shoots dry weight (g) of *O. sativa* at 6, 9 and 12 weeks following inoculation. Black bars represent the growth of plants inoculated with parental lines of AMF *Glomus intraradices*; white bars represent the growth of plants inoculated with crossed lines resulting from cross between C3 and D1 lines; hatched bars represent the growth of the uninoculated treatment. Errors bars Error bars ± 1SE.

Genetic exchange significantly altered the AMF colonization levels inside rice roots (Figure 2, Table 3). Differences in colonization levels were large. Overall, the crossed lines Sb and Sd exhibited lower colonization levels compared to the other crossed line (Sc2) and compared to the two parental lines (Figure 2; Table 3). Interestingly, the differences in the fungal colonization among AMF lines were not the same through time (shown by a significant line by harvest time interaction; Table 3). For example, at 6 weeks after inoculation, parental line D1 showed the highest hyphal colonization while at 9 and 12 weeks after inoculation, no differences were found between the 2 parental lines and the crossed line Sc2 (Figure 2).

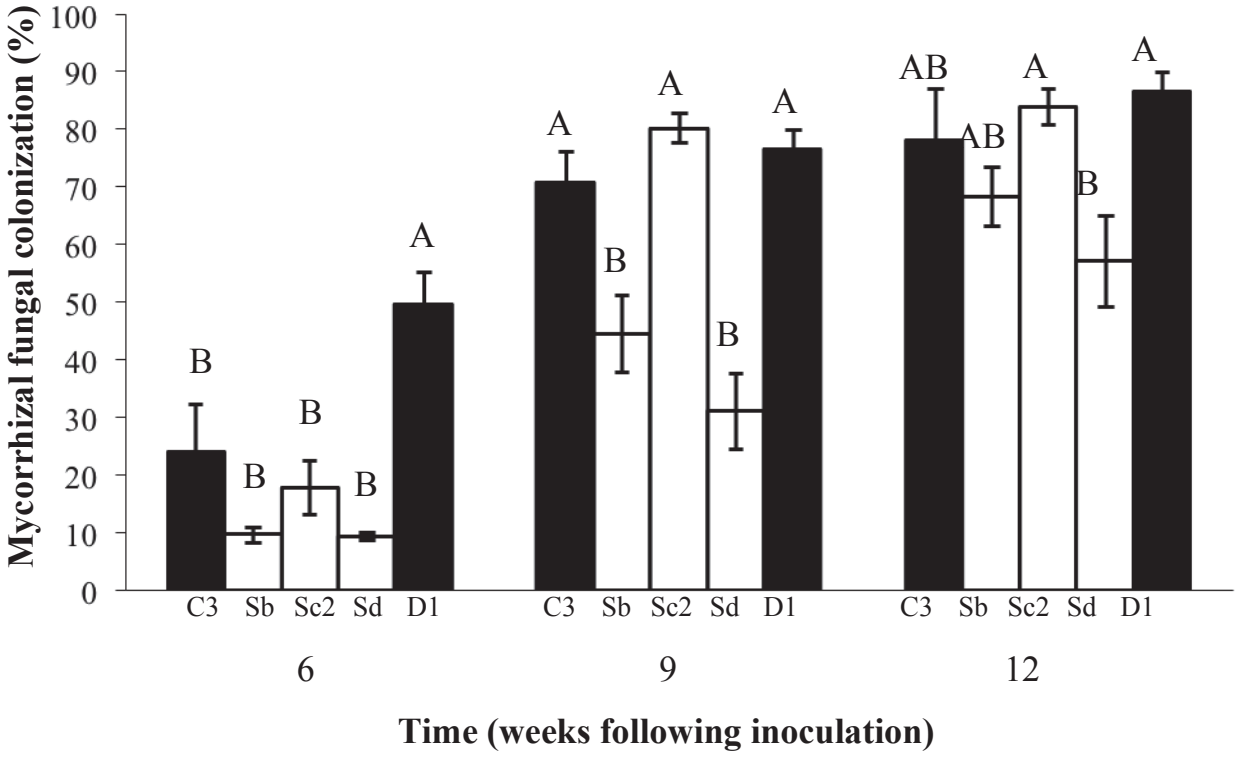


Figure 2: Mycorrhizal fungal colonization (% root length colonized) in the roots of *O. sativa* at 6, 9 and 12 weeks after inoculation. Black bars represent colonization by parental lines and white bars by crossed lines, originating from crossed between C3 and D1. Error bars ± 1SE. Bars topped by different letters within each harvest are significantly different at $p \leq 0.05$.

Table 3: Analysis of variance on plant shoot dry weight and fungal colonization.

Source	d.f.	SS	F	*p*
Plant shoot dry weight				
Harvest time	**2**	**0.32**	**66.55**	**<0.0001*** **
AMF line	**4**	**0.03**	**2.86**	**0.027***
Harvest time x AMF line	**8**	**0.02**	**0.83**	**0.58**
Fungal colonization				
Harvest time	**2**	**5.71**	**94.68**	**<0.0001*** **
AMF line	**4**	**2.53**	**20.96**	**<0.0001*** **
Harvest time x AMF line	**8**	**0.61**	**2.54**	**0.014***

Mycorrhizal specific gene transcription

Differences in plant gene transcription among AMF lines were studied for each harvest time separately. Results showed that genetic exchange in AMF had an effect on mycorrhizal specific plant gene transcription (Figure 3; Table 4). We found significant differences in the transcription of AM1 and AM3 in rice inoculated with different AMF lines 6 weeks following inoculation. Differences in transcription occurred both between plants inoculated with parental and crossed lines and also between plants inoculated with different crossed lines (Figure 3a, and 3b). There was also a significant difference in the transcription of AM10 and AM14 at 12 weeks after inoculation between rice inoculated with parental lines and also between rice inoculated with parental and some crossed lines (Figure 3c and 3d, Table 4).

Correlations between gene transcription, fungal colonization and plant growth

We found significant positive correlations between gene transcription and fungal colonization at 6 weeks after inoculation for the AM1 ($r = 0.90$; $p = 0.003$), AM14 ($r = 0.94$; $p = 0.003$) and Pt11 ($r = 0.90$; $p = 0.003$) genes. AM11 gene transcription was negatively correlated with shoot dry weight ($r = -0.88$; $p = 0.048$), but positively correlated with the arbuscular colonization ($r = 0.88$; $p = 0.049$) at 9 weeks after inoculation. Another negative correlation

was found between transcription of another gene, AM25, and hyphal colonization (r = -0.889; p = 0.0434). All the other genes (except AM29) showed the same trend as AM25 at 12 weeks after inoculation but correlations were not significant. We also found a positive correlation between arbuscular colonization and shoot dry weight (r = 0.399; p = 0.0074) at 12 weeks after inoculation.

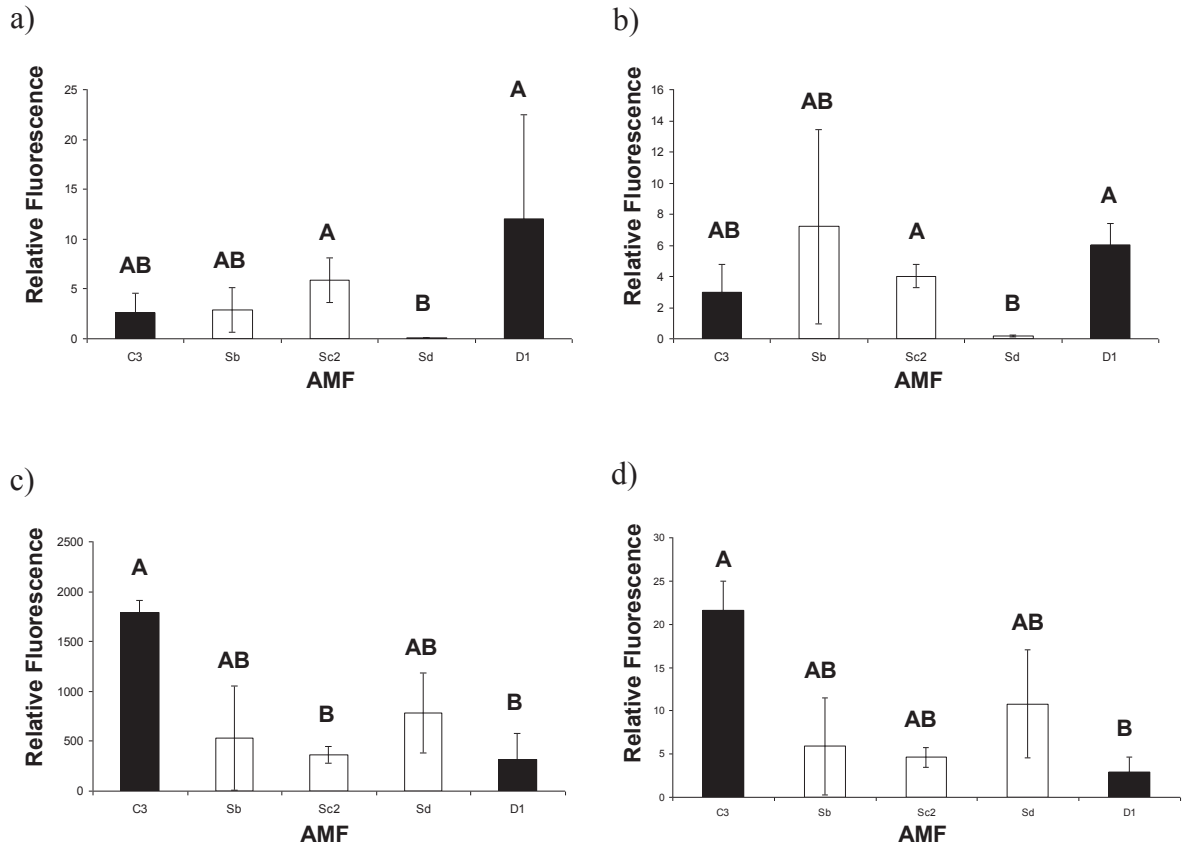


Figure 3: Effect of genetic exchange in AMF on transcription of AM1 (a), AM3 (b), AM10 (c) and AM14 (d) genes in the roots of *O.sativa*. Black bars represent the relative transcription of genes when plants were inoculated with parental lines C3 and D1. White bars represent the relative transcription of genes when plants were inoculated with crossed lines Sb, Sc2 and Sd originating from crosses between C3 and D1 lines. Error bars ± 1 SE. Bars topped by different letters are significantly different at $p \leq 0.05$.

Table 4: Analysis of variance on plant gene transcription at 6 and 12 weeks after inoculation for the genes AM1, AM3, AM10 and AM14.

Source	d.f.	SS	F	*p*
6 weeks after inoculation				
AM1	**4**	**72.36**	**4.90**	**0.027***
AM3	**4**	**81.25**	**4.00**	**0.045***
12 weeks after inoculation				
AM10	**4**	**4338952.1**	**4.78**	**0.024***
AM14	**4**	**674.01**	**3.74**	**0.047***

Discussion

Here, we showed that genetic exchange in the AMF *Glomus intraradices* alters rice growth and gene transcription. Gutjahr *et al*. (2008) characterized genes that are only transcribed in rice roots during the establishment of the mycorrhizal symbiosis. We used 12 of these genes and measured their transcription at 3, 6, 9 and 12 weeks after inoculation with two genetically different AMF parental lines and three AMF crossed lines arosing from genetic exchange between the parental lines. Our results support our hypothesis that genetic exchange in AMF alters the transcription of some of these genes compared to the transcription in plants inoculated with parental lines. We also demonstrated that genetic exchange in AMF can alter the speed of AMF colonization and also increase the growth of rice. Finally, we found that some of the transcriptional and growth changes were correlated with colonization, but not all.

Effect of genetic exchange on rice growth and AMF colonization

Angelard *et al.* (2010) showed that segregation can improve rice growth by a factor five. They also tested the effect of genetic exchange in AMF on rice growth. They found that genetic exchange in AMF did not improve the plant growth. Indeed, they observed a better growth of uninoculated plants and plants inoculated with parental lines than plants inoculated with crossed lines. Furthermore, in those experiments rice always grew significantly larger when non-mycorrhizal compared to any of the mycorrhizal treatments. In contrast, we found an overall positive growth effect of the AMF lines on rice through time, with some crossed lines inducing an even higher plant growth than the parental lines throughout the experiment. We also found that crossed and parental lines did not colonize rice roots in a similar way over time. Angelard and Sanders (2011) also measured the fungal colonization inside rice roots. They found that the crossed lines Sb and Sc2 exhibited a lower colonization in rice than the other AMF lines while we found that the crossed line Sd exhibited the lowest colonization. The environmental conditions were markedly different between the experiments of Angelard *et al.*(2010) and our experiment. This could explain the different results on plant growth and fungal colonization. Our experiment was performed in a phytochamber with strictly controlled environment and the experiment of Angelard *et al.* (2010) was performed in greenhouse where the environment could have varied greatly during the experiment in terms

of light intensity, humidity and temperature. Furthermore, the pot size in this experiment was much smaller than that used by Angelard *et al.* (2010). The contrasting results obtained from the two experiments showed that the effects of AMF crossed lines on plants are likely to be dependent upon the environment. It also shows the large potential range of variation in the effects of an AMF and this should be taken into account for further laboratory or field experiments.

Plant gene expression

Symbiosis-specific rice genes were found after an extensive study in 2005 (Guimil *et al.* 2005). 224 genes were shown to only be transcribed when the mycorrhizal fungus *Glomus intraradices* was present in the soil. A recent study characterized 18 of these genes (Gutjahr *et al.* 2008) and measured their transcription at 3, 5, 7 and 9 weeks after inoculation. Some of them played a role in early transcription (less than 7 weeks after inoculation) and others were characterized as lately transcribed (at least 7 weeks after inoculation). Angelard *et al.* (Angelard *et al.* 2010) looked at 4 of these genes (2 "early" (AM1 and AM3) and 2 "late" transcribed genes (AM14 and Pt11) and showed that segregation in AMF had an effect on rice AM-specific gene transcription. In our study, we used the same four genes, plus eight other AM-specific genes (Table 2) identified by Gutjahr *et al.* (2008). They are classified into 3 "early" genes and 9 "late" genes. We found a significant effect of AMF lines on the transcription of 4 genes: AM1, AM3, AM10 and AM14 (Figure 3, Table 4). Interestingly, the time at which we found significant effects corresponds to whether the genes were characterized as early or late transcribed. Only putative functions are known for these genes (Table 2). AM1 is a putative peroxidase, AM3 contains a peptidoglycan binding LysM domain, AM14 is a serine-threonine kinase-like gene and AM10 is similar to hypersensitivity–related (Hsr) protein. Angelard *et al.* (2010) found that segregation has significant effects on AM1, AM3, AM14 and Pt11. It is interesting to note that we also found significant effects for 3 of these genes. Even if the functions of these genes in the symbiosis are not known, we show here that transcription of those genes is altered by the genetic identity of the fungus following genetic exchange between two genetically distinct AMF.

We showed that symbiosis-specific gene expression was positively correlated with fungal colonization for some of the genes but not all. In our experiments, genetic exchange in AMF had strong effects on colonization rates. It could therefore have been possible that the changes

in symbiosis-specific gene expression would simply reflect the speed of a given crossed line to colonize rice roots. However, the lack of significant correlation between transcription of some of the affected genes and colonization levels suggests that the effect on transcription is not so simple.

Conclusions

In view of the results of our study, further experiments using microarrays or high-throughput sequencing approaches are highly warranted as they could give a better understanding of how genetic changes in mycorrhizal fungi affect patterns of rice gene transcription over the whole transcriptome.

Angelard *et al.* (2010) already showed that segregation in the AMF *Glomus intraradices* can improve growth of rice up to five times. Our study shows, for the first time, that genetic exchange in AMF can also have beneficial effects on rice growth. Rice is probably the most important cereal on earth. It feeds billions of people every day. Improving rice productivity is a priority with the growth of the world population. Consequently, these recent advances in our knowledge of the contribution of AMF genetics to plant productivity could potentially be used for improving crop productivity by manipulating the fungus genetically, using the fungi's own natural genetic processes.

Acknowledgments

This work was supported by grants 31000AO-105790/1 and 31003A-127371 from the Swiss National Science Foundation. We thank C. Gutjahr for assistance in the quantification of gene transcripts in rice, U. Paszkowski for providing rice seeds, J. Bonvin for assistance with fungal inoculation, Daniel Croll for providing the crossed lines and the rest of our group for their useful comments on the manuscript. We also thank the Genomic Technology Facility of the University of Lausanne for their assistance for the quantitative PCR.

Chapter 2

Using natural genetic variation in arbuscular mycorrhizal fungi as an additional approach to identifying genes involved in the mycorrhizal symbiosis

Alexandre Colard, Caroline Angelard, Jose Joaquim Marulanda Martinez, and Ian R. Sanders

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

Abstract

Arbuscular mycorrhizal fungi (AMF) form symbiosis with most land plants. The plants supply AMF with photosynthetically fixed carbon. In exchange, AMF increase nutrients uptake from soil, particularly phosphate and are thus, able to improve plant growth. Until recently, only few studies have investigated the molecular basis of the symbiosis. These studies shown by inoculated/ non-inoculated comparison that some plant genes are mycorrhizal specific. Because others studies have shown that genetically different AMF coming from the same population could have different effects on plant growth, we investigated the effect of such AMF isolates on plant transcriptome. We found that pathways involved in carbohydrates were up-regulated and those involved in plant defences were down-regulated in inoculated plants compared with non-inoculated plants. Interestingly, no pathways involved in phosphate transport were altered. We also showed that some pathways were specific depending of the inoculated isolate. By comparison of the transcriptome profile of plants inoculated with different AMF isolates, we found that some chitinases transcriptions were altered as well as pathways involved in transport of oxygen. In overall, results showed that genetically different AMF can have a large effect on plant transcriptome profile and whether different transcription profiles would be generated among plants inoculated with two genetically different AMF isolates of the same species and non-mycorrhyzal plants. This approach allowed us to identifying new genes that were sensitive to the AM symbiosis.

Introduction

The arbuscular mycorrhizal (AM) symbiosis is probably the most widespread symbiosis in terrestrial ecosystems. It occurs between plant roots and arbuscular mycorrhizal fungi (AMF)(Smith *et al.* 2008). Plants receive phosphate (Smith *et al.* 2008), nitrogen and other nutrients from the fungi (Hodge *et al.* 2001). AMF also make plants more resistant to different stresses, such as heavy metal stresses (Ricken *et al.* 1996), herbivory (Gange *et al.* 2003) and drought (Ebel *et al.* 1996). In exchange, the plant supplies carbohydrates to the fungus (Jakobsen *et al.* 1990). Studies have also shown the importance of AMF diversity for ecosystem productivity and plant diversity (van der Heijden *et al.* 1998; van der Heijden *et al.* 1998; Klironomos *et al.* 2000). All the major crops in the world form the mycorrhizal symbiosis and given the effects of the symbiosis on plant productivity, the symbiosis should be of major importance in agriculture.

To better understand the functioning of the symbiosis, researchers are engaged in trying to identify the plant and fungal genes involved both in the establishment and development of the symbiosis, as well as all the genes that are up- and down-regulated by the symbiosis (Harrison *et al.* 2002; Requena *et al.* 2002; Liu *et al.* 2004; Weidmann *et al.* 2004; Massoumou *et al.* 2007; Parniske 2008; Feddermann *et al.* 2010). Such investigations have mainly been conducted using two traditional approaches. First, a comparison between wild type plants and plant mutants that do not fully develop a normal symbiosis and the identification of genes involved in normal development of the symbiosis (Reddy *et al.* 2007; Feddermann *et al.* 2010). Second, the determination of differentially transcribed or expressed genes between mycorrhizal and non-mycorrhizal plants (Guimil *et al.* 2005; Hohnjec *et al.* 2005; Fiorilli *et al.* 2009; Breuillin *et al.* 2010). Guimil *et al.* (2005) carried out a large scale study where they compared microarray data of gene transcription of AMF inoculated and uninoculated rice. The authors analysed 48564 genes and showed that 224 of them had a differential transcription levels according to whether they were mycorrhizal or non-mycorrhizal. Among the 224 genes, 129 were only transcribed in the presence of AMF and were, therefore, considered as AM symbiosis specific genes. Later, a study characterized the transcription of 38 of these genes over time and showed the localization of some (Gutjahr *et al.* 2008). Most of these genes only have a putative function, except Pt11 that was already characterized as a phosphate transporter (Paszkowski *et al.* 2002)

However, over the last decade it has become apparent that the genetic identity of the fungal partner also differentially affects plant growth. Several studies showed that plants grow differently when inoculated with different AMF species (van der Heijden *et al.* 1998; van der Heijden *et al.* 2003). More recently, it has been shown that genetically different AMF of the same species also differentially alter plant growth and inorganic phosphate acquisition (Munkvold *et al.* 2004; Koch *et al.* 2006). The strong potential role of AMF genetics in the plant response to the symbiosis is highlighted by the recent finding that intra-specific genetic diversity in AMF can lead to changes in plant growth and plant gene expression (Angelard *et al.* 2011; Colard *et al.*, *chapter1*). Interestingly, in those studies, genetic differences among the fungi induced large differences in transcript levels of genes that had been identified as mycorrhizal-specific in traditional mycorrhizal versus non-mycorrhizal studies.

We hypothesized that because genetically different AMF alter plant growth and also the plants mycorrhizal specific gene transcription, other plant genes must be involved in the symbiosis that would not necessarily be detected in a classic mycorrhizal versus non-mycorrhizal comparison. In order to test this, we chose two genetically different lines of the AMF species *Glomus intraradices* coming from the same population and that have already been shown to alter plant growth. *Oryza sativa* was inoculated with these two AMF or left uninoculated. 12 weeks after inoculation, total RNA was extracted from the roots and then hybridized with rice microarrays carrying 48564 genes (55509 probe sets). In addition to the identification of transcripts altered by the presence of the symbiotic partner, this study allows us to detect variation in transcript levels caused by the genetic variation between the two isolates.

Material and methods

Origin of the fungal lines and plants

Two *Glomus intraradices* lines, C2 and D1, were used in this study. They originate from one agricultural field in Tänikon, Switzerland (Koch *et al.* 2004). The fungi were isolated from two different plots in the field, each separated by 60 meters. These two AMF lines were previously shown to be genetically distinct, to differ in their phenotypic traits (Koch *et al.* 2004; Croll *et al.* 2009) and to differentially affect plant growth (Koch *et al.* 2006; Angelard *et al.* 2010). The two lines were cultivated in identical conditions with Ri T-DNA transformed carrot roots (Becard *et al.* 1988) and maintained through clonal subculturing with the same protocol as in Koch *et al*. (2004).. We used *Oryza sativa* (Nipponbare) for these experiments because this variety had previously been used for studies of AM induced transcriptional changes.

Plant inoculation and harvesting

AMF spores were extracted from four month old cultures using the same method as in Koch *et al.* (2006). The plant inoculation protocol was the same as that performed by Angelard *et al.* (2010). Each plant was inoculated with 500 spores of one of the AMF lines. Some plants were not inoculated and served as the non-mycorrhizal control. There were ten replicates per inoculated treatment, giving a total of 20 pots plus 10 pots for the non-mycorrhizal treatment. The position of the pots in the greenhouse was randomized. Each pot was watered with 50-75 ml of tap water every two days. Plants were fertilized twice during the experiment with a full strength Hoaglands solution containing no phosphorous. Potassium concentration was adjusted by adding KCl. Growth conditions in the greenhouse were 16 hours daylight, and temperatures vary between 18°C and 30°C. Under these conditions, *O. sativa* of the Nipponbare variety are expected to go through the following developmental stages: germination, seedling establishment, tillering, stem elongation, booting or panicle initiation and heading (Yamamoto *et al.* 2000).

After 12 weeks, the shoots of each plant were harvested, dried at 80°C for two days and weighed. The roots were immediately washed and separated into two parts. Two thirds of the roots were frozen in liquid nitrogen for subsequent RNA extraction and stored at -80°C. The remaining one third was stored in 50% ethanol and then stained with trypan blue as described in Munkvold *et al.* (2004) in order to check for fungal colonization. Plants roots were checked for presence or absence of AMF colonization but the percentage of colonization was not

measured. All the non-mycorrhizal (NM) plants were found to be AMF free. All plants that were inoculated were colonized except for one plant. One plant died during the experiment. Those two plants were removed from the data, which left 28 individuals for the analyses.

RNA extraction

RNA extractions were performed directly on frozen roots. Roots were ground with a pestle in liquid nitrogen. The RNeasy Plant mini kit (Qiagen Inc®) was used to extract RNA following the manufacturer's recommendations. Precipitation with LiCl was also performed (Ambion®) to remove sugar present in plant roots. DNase treatment was performed with DNase I grad (Invitrogen®) following the manufacturer's recommendations. A PCR on extracted RNA was performed to amplify the housekeeping genes cyclophilin and ubiquitin to be sure that there was no DNA contamination, with the following conditions: 94°C for 5 minutes, 35 cycles [94°C 20 seconds, 58°C 20 seconds, 72°C 50 seconds] 72°C for 7 minutes.

Microarray hybridization

RNA amplification and GeneChip rice genome array hybridization (also known as affimetrix 57) were carried out at the microarray facility of the Center for Integrative Genomics at the University of Lausanne according to the following protocol. Because a large quantity of RNA was needed, WT-Ovation™ Pico RNA Amplification System from NuGEN™ was used following the manufacturer's recommendations to amplify plant RNA and transcript it into cDNA. Labelling of the cDNA was done with the FL-Ovation™ cDNA Biotin Module V2 from NuGEN™ following the manufacturer's recommendations. At this step, three samples of each treatment were merged to form three pooled replicates per treatment. One chip was used to hybridize with each of the pooled replicate samples in each treatment, making a total of nine chips in the complete experiment. Affymetrix Rice Genome arrays (Affymetrix, Santa Clara , CA, USA) were hybridized with 5 microgrammes (μg) of biotinylated target, at 45°C for 17 h washed and stained according to the protocol described in Affymetrix GeneChip® Expression Analysis Manual (Fluidics protocol FS450_0004). The arrays were scanned using the GeneChip® Scanner 3000 7G (Affymetrix).

Quality control of microarray hybridization

After image acquisition, the scripts Affy (Gautier *et al.* 2004), AffyPLM (Irizarry *et al.* 2003) and Limma (Smyth 2005) implemented in bioconductor software (Gentleman *et al.* 2004), R language (Team 2008), were used to extract the information from the raw data files. No potential artifacts or background gradients were detected. Since no technical problems were found, data from all the chips used for hybridization in our experiment were further analyzed.

Using customized scripts developed for this project (available upon request) boxplots for unnormalized data were printed as a first step in the hybridization quality control. Additionally, the normalized unscaled standard error (NUSE) procedure was used as a quality control measure. Raw and background intensity images (probe level model residuals (PLM)) of the chips were printed to check for quality and spatial heterogeneity inside and between the chips. Normalization was applied via the Robust Multi-Array method (RMA) which uses a convolution background correction, quantile normalization and summarization (median polish) using a robust multi-array linear model. The efficiency of the normalization procedure, revealed in the boxplots, allows further comparisons using linear models and t-tests, which are based on a normal distribution of the data. After normalization, expression files were extracted and printed to be used as input data for further analysis.

Gene Ontology Analysis

Probe sets were selected based on their fold change (FC). Probe sets were considered to be affected by treatment when FC was upper than two. Usual approaches use the Benjamini-Hochberg False Discovery Rate (FDR) statistic to detect affected probe sets. Because the two isolates C2 and D1 are originated from the same population, almost no FDR values coming from the basic analysis from Limma package was significant. Therefore, we based level changes on FC as a signal of presence/absence. We selected the threshold of 2 because the low number of probe sets with threshold greater than 3 or 4 did not permit us to perform the following analysis.

The lists of probe sets affected by the treatments were used to carry out a singular enrichment analysis (SEA) in AgriGO, a web-based tool and database to perform Gene Ontology (GO) analysis, especially for agriculturally relevant species (Du *et al.* 2010). GO is used widely in functional annotation and enrichment analysis. The enrichment analysis classifies the probe sets in the different GO terms, and gives as output the pathways that are significantly over represented in the microarray data. GO project is a major bioinformatics initiative with the aim

of standardizing the representation of genes and gene product attributes across many different species and databases. The well-annotated Gramene database (rice gramene locus) was used as a background for our lists corresponding to *Oryza sativa* (var. japonica). At the same time, a fisher statistical test method with a Benjamini-Hochberg FDR correction was implemented. The significance level was set up at 0.05 and the minimum number of mapping entries was five. The fact to retrieved enrich pathways with our different group of probe sets is a clue to attest the reliability of the microarray output.

Microarray validation

We selected probe sets that were up- or down-regulated, and we carried out quantitative PCR on each of them. Quantitative PCR results were congruent with microarray results (see supplementary file Figure S1 and Table S1).

Results

Effect of isolates C2 and D1 on plant growth

Twelve weeks after germination rice plants were harvested. There was no significant effect of AMF inoculation on dry weight of rice (data not shown). Colonization level was checked for each plant roots. Levels of colonization were not measure but plant roots used in the follow analysis were well colonized.

Overall effects of the inoculation treatments on rice gene transcription

As described in material and methods, we used probe sets with FC>2 values as a threshold. We were able to detect 2219 probe sets that were differentially transcribed between plants inoculated with *G. intraradices* and plants that were uninoculated (Figure 1). With thresholds of FC>3 and FC>4, we obtained 104 probe sets and 39 probe sets that were differentially transcribed, respectively. With the same threshold of FC>2, we detected 1089 probe sets that were differentially transcribed in plants inoculated with the C2 and D1 isolates (Figure 1). With thresholds of FC>3 and FC>4, we obtained 241 probe sets and 112 probe sets that were differentially transcribed, respectively.

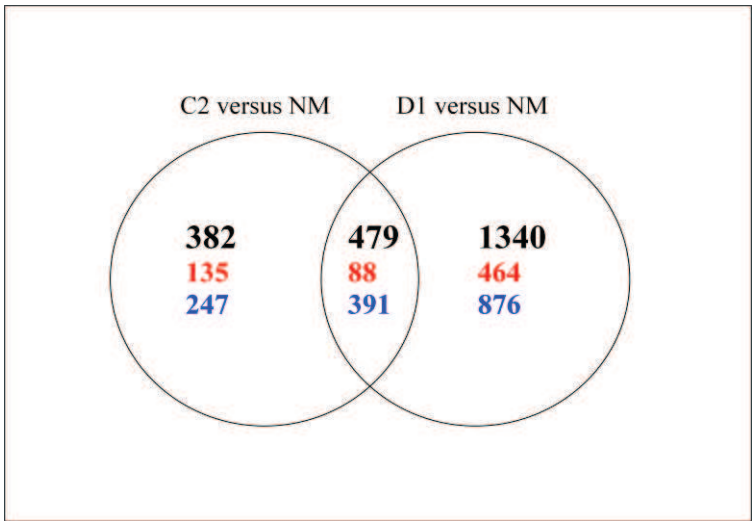


Figure 1: Blue numbers represent number of probe set that are down regulated; Red numbers represent number of probe set that are up regulated; Black numbers represent the total number of probe set.

Probe sets commonly up or down-regulated by inoculation with G.intraradices *compared to non-mycorrhizal plants*

Among the 2219 probe sets that were differentially transcribed between mycorrhizal plants and non-mycorrhizal plants, we detected 497 probe sets that were common to plants inoculated with either C2 or D1 isolates (Figure 1). The majority of these (461 probe sets) were affected in the same way in plants inoculated with C2 and D1 isolates compared to the non-mycorrhizal treatment. Only, three probe sets were up-regulated in plants inoculated with D1 isolate that were down-regulated in plants inoculated with C2 isolate.

GO	Type	Functions	Count	p-value	FDR
GO:0009250	BP	glucan biosynthetic process	6	1.30E-09	**1.90E-07**
GO:0033692	BP	cellular polysaccharide biosynthetic process	6	4.40E-09	**3.40E-07**
GO:0000271	BP	polysaccharide biosynthetic process	6	1.10E-08	**5.60E-07**
GO:0016137	BP	glycoside metabolic process	6	5.10E-07	**1.20E-05**
GO:0009311	BP	oligosaccharide metabolic process	6	5.30E-07	**1.20E-05**
GO:0005984	BP	disaccharide metabolic process	6	4.90E-07	**1.20E-05**
GO:0005985	BP	sucrose metabolic process	6	4.40E-07	**1.20E-05**
GO:0005982	BP	starch metabolic process	6	6.20E-07	**1.20E-05**
GO:0006073	BP	cellular glucan metabolic process	6	1.60E-06	**2.50E-05**
GO:0044042	BP	glucan metabolic process	6	1.60E-06	**2.50E-05**
GO:0034637	BP	cellular carbohydrate biosynthetic process	6	1.90E-06	**2.70E-05**
GO:0044264	BP	cellular polysaccharide metabolic process	6	2.50E-06	**3.20E-05**
GO:0016051	BP	carbohydrate biosynthetic process	6	8.10E-06	**9.40E-05**
GO:0005976	BP	polysaccharide metabolic process	6	8.40E-06	**9.40E-05**
GO:0006006	BP	glucose metabolic process	5	1.60E-05	**0.00016**
GO:0005975	BP	carbohydrate metabolic process	8	5.30E-05	**0.00051**
GO:0019318	BP	hexose metabolic process	5	6.40E-05	**0.00058**
GO:0044262	BP	cellular carbohydrate metabolic process	6	0.00013	**0.001**
GO:0005996	BP	monosaccharide metabolic process	5	0.00013	**0.001**
GO:0034641	BP	cellular nitrogen compound metabolic process	6	0.00023	**0.0018**
GO:0006066	BP	alcohol metabolic process	5	0.0003	**0.0022**
GO:0009308	BP	amine metabolic process	5	0.036	0.26
GO:0016740	MF	transferase activity	15	0.0029	**0.062**
GO:0043169	MF	cation binding	11	0.0094	**0.068**
GO:0043167	MF	ion binding	11	0.0094	**0.068**
GO:0046872	MF	metal ion binding	9	0.04	0.18
GO:0003824	MF	catalytic activity	23	0.041	0.18
GO:0009507	CC	chloroplast	6	0.0048	0.3

Table 1: significant GO terms for genes commonly up-regulated in plants inoculated with C2 and D1 isolates compared to the uninoculated treatment.

GO: Reference number of GO term, Type: BP: Biological processes; MF: Molecular functions; CC: Cell cycle. Count is the number of up-regulated genes found in a given GO term. P-value is given for Fisher test. FDR: False discovery Rate. In bold are present the significant values (≥0.1 for the FDR).

Genes corresponding to these probe sets encode for unknown proteins, but homologous genes have been detected in *Medicago truncatula* (Table S2). Similarly, 15 probe sets were found to be up-regulated in plants inoculated with isolate C2, and down-regulated in plants inoculated with isolate D1. These probe sets belong to different protein encoding families such as nitrate reductase, glutamine synthestase or ferredoxin-6. All these genes and their putative functions are listed in Table S3. However, Gene Ontology (GO) analysis did not permit us to identify any specific pathways concerning these 15 probe sets.

Eighty-eight probe sets were commonly up-regulated in mycorrhizal plants compared to non-mycorrhizal plants (Figure 1). The GO analysis allowed us to detect probe sets that were common to specific pathways (Table 1). For example, pathways involved in carbohydrate metabolism and in different types of sugar metabolism were up-regulated in mycorrhizal plants compared to non-mycorrhizal plant. We also observed that pathways involved in carbon storage were up regulated.

Three hundred ninety one probe sets were found to be down-regulated in mycorrhizal plants compared to non-mycorrhizal plants (Figure 1). Results of the GO analysis revealed that most of the pathways that were down-regulated were involved in chitinase activity, hydrolase activity, responses to stress, responses to other organisms and defence responses (Table 2). A summary of all these results is given by a heat map in Figure 2.

Probe sets exclusively regulated in plants inoculated with G. intraradices (isolate C2).

Three hundred and eighty-two probe sets were differentially transcribed in plants inoculated exclusively with the AMF isolate C2 compared to the non-mycorrhizal plants (Figure 1). Among these 382 probe sets, 247 probe sets were down-regulated and 135 probe sets were up-regulated. The GO analysis allowed us to identify pathways specifically affected by the isolate C2. Down-regulated probe sets included those in pathways concerning response to other organisms, response to biotic stimulus, response to stimulus and response to stress and defence response (Table 3). Moreover, some hydrolase activities were down-regulated as well as one carbohydrate metabolism pathway. GO analysis on the 135 up-regulated probe sets revealed that some GO terms were up-regulated, including one pathway concerning transport (Table 4). This GO term (GO: 0006810) is composed of 117 genes. Nine of the 135 up-regulated probe sets are present in this GO term. Most of these probe sets are involved in sugar transport (Table S4). One of them is known to be a Glucose-6-phosphate/phosphate translocator, involved in exchange of carbon for inorganic phosphate.

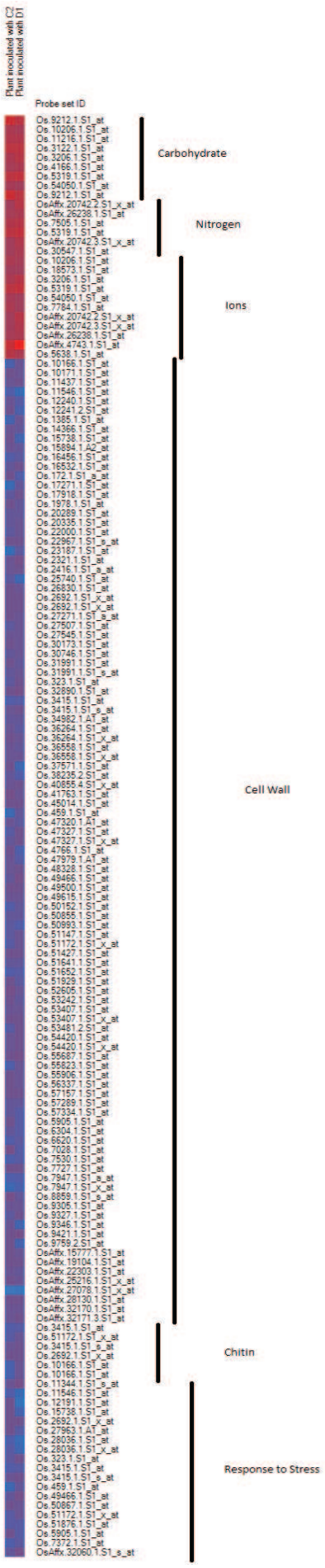


Figure 2: Heat map of the differentially transcribed probe sets comparing plants inoculated with C2 (first column) or D1 (second column) versus plants in the uninoculated treatment. Each probe set is clustered by metabolic pathways.

Each square represents the transcription level of each probe set. Colour of squares represents the transcription level of the probe sets. Red squares represent probe sets with a transcription higher than in the uninoculated treatment. Blue squares represented probe sets with transcription levels below those in the uninoculated treatment.

GO	Type	Functions	Count	p-value	FDR
GO:0016998	BP	cell wall macromolecule catabolic process	7	2.00E-09	**4.60E-07**
GO:0006032	BP	chitin catabolic process	5	2.60E-08	**3.00E-06**
GO:0044036	BP	cell wall macromolecule metabolic process	7	8.70E-08	**5.60E-06**
GO:0006030	BP	chitin metabolic process	5	1.20E-07	**5.60E-06**
GO:0006026	BP	aminoglycan catabolic process	5	1.20E-07	**5.60E-06**
GO:0006022	BP	aminoglycan metabolic process	5	3.00E-06	**0.00012**
GO:0009607	BP	response to biotic stimulus	8	5.10E-05	**0.0017**
GO:0000272	BP	polysaccharide catabolic process	5	7.00E-05	**0.002**
GO:0051707	BP	response to other organism	7	0.00019	**0.0049**
GO:0006950	BP	response to stress	16	0.0011	**0.026**
GO:0051704	BP	multi-organism process	7	0.0016	**0.033**
GO:0050896	BP	response to stimulus	22	0.0017	**0.033**
GO:0006952	BP	defense response	6	0.01	0.18
GO:0016052	BP	carbohydrate catabolic process	5	0.017	0.28
GO:0005976	BP	polysaccharide metabolic process	5	0.039	0.6
GO:0008061	MF	chitin binding	5	1.20E-09	**1.20E-07**
GO:0030247	MF	polysaccharide binding	5	8.60E-08	**3.00E-06**
GO:0001871	MF	pattern binding	5	8.60E-08	**3.00E-06**
GO:0020037	MF	heme binding	10	3.70E-06	**9.60E-05**
GO:0046906	MF	tetrapyrrole binding	10	6.00E-06	**0.00013**
GO:0004568	MF	chitinase activity	5	2.80E-05	**0.00048**
GO:0030246	MF	carbohydrate binding	5	0.0015	**0.018**
GO:0005506	MF	iron ion binding	10	0.0016	**0.018**
GO:0004553	MF	hydrolase activity	7	0.0013	**0.018**
GO:0016798	MF	hydrolase activity	7	0.0033	**0.035**
GO:0043167	MF	ion binding	28	0.031	0.25
GO:0043169	MF	cation binding	28	0.031	0.25
GO:0046914	MF	transition metal ion binding	21	0.031	0.25
GO:0005215	MF	transporter activity	12	0.043	0.32
GO:0031982	CC	vesicle	100	2.10E-07	**4.10E-06**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	100	2.00E-07	**4.10E-06**
GO:0031988	CC	membrane-bounded vesicle	100	2.00E-07	**4.10E-06**
GO:0031410	CC	cytoplasmic vesicle	100	2.10E-07	**4.10E-06**
GO:0043231	CC	intracellular membrane-bounded organelle	221	0.031	0.32
GO:0043227	CC	membrane-bounded organelle	221	0.031	0.32
GO:0044444	CC	cytoplasmic part	214	0.024	0.32
GO:0005737	CC	cytoplasm	216	0.034	0.32
GO:0043229	CC	intracellular organelle	221	0.048	0.37
GO:0043226	CC	organelle	221	0.048	0.37

Table 2: Significant GO terms for genes commonly down-regulated in plants inoculated with C2 and D1 isolates compared to the uninoculated treatment. Column headings follow Table1.

Probe sets exclusively regulated in plants inoculated with G. intraradices (isolate D1).

The transcription of 1340 probe sets was exclusively altered in plants inoculated with D1 isolate compared to the non-mycorrhizal plants. 464 probe sets were up-regulated and 876 were down-regulated (Figure 1). Results of the GO analysis on the 464 up-regulated probe sets revealed that the pathways altered were almost entirely related to functions within organelles (Table S5). GO analysis on the 876 down-regulated probe sets shows that a very large number of different pathways were altered (Table S6), many with a large number of probe sets per GO terms.

GO	Type	Functions	Count	p-value	FDR
GO:0051707	BP	response to other organism	6	0.00016	**0.025**
GO:0009607	BP	response to biotic stimulus	6	0.00026	**0.025**
GO:0050896	BP	response to stimulus	19	0.00023	**0.025**
GO:0006950	BP	response to stress	13	0.00068	**0.049**
GO:0005975	BP	carbohydrate metabolic process	11	0.00087	**0.051**
GO:0051704	BP	multi-organism process	6	0.0011	**0.054**
GO:0006952	BP	defense response	6	0.0016	**0.065**
GO:0005976	BP	polysaccharide metabolic process	6	0.0018	**0.065**
GO:0004553	MF	hydrolase activity	8	1.70E-05	**0.0024**
GO:0016798	MF	bonds	8	5.90E-05	**0.0041**
GO:0003824	MF	catalytic activity	63	0.0021	**0.064**
GO:0016491	MF	oxidoreductase activity	14	0.0023	**0.064**
GO:0005506	MF	iron ion binding	8	0.0017	**0.064**
GO:0020037	MF	heme binding	5	0.0038	**0.089**
GO:0046906	MF	tetrapyrrole binding	5	0.0048	**0.096**
GO:0043169	MF	cation binding	20	0.039	0.62
GO:0043167	MF	ion binding	20	0.04	0.62
GO:0031982	CC	vesicle	62	0.00028	**0.0065**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	62	0.00028	**0.0065**
GO:0031988	CC	membrane-bounded vesicle	62	0.00028	**0.0065**
GO:0031410	CC	cytoplasmic vesicle	62	0.00028	**0.0065**

Table 3: GO terms for genes significantly down-regulated exclusively in plants inoculated with C2 isolate compared to the uninoculated treatment. Column headings follow Table1.

Differences observed in transcription profiles between plants inoculated with C2 or D1 isolates

A comparison between the transcription profiles coming from plants inoculated with C2 isolate and D1 isolate was performed. Transcription was affected in one thousand and eighty nine probe sets, with the same threshold used previously. Among them, 413 were found to be up-regulated in plants inoculated with isolate D1 in comparison with plants inoculated with isolate C2, and 676 were found to be down-regulated (Figure 3).

GO	Type	Functions	Count	p-value	FDR
GO:0006575	BP	cellular amino acid derivative metabolic process	5	0.0019	0.2
GO:0006810	BP	transport	9	0.031	0.88
GO:0051234	BP	establishment of localization	9	0.031	0.88
GO:0051179	BP	localization	9	0.034	0.88
GO:0005215	MF	transporter activity	9	0.001	**0.039**
GO:0016491	MF	oxidoreductase activity	7	0.035	0.64
GO:0031988	CC	membrane-bounded vesicle	29	0.036	0.44
GO:0031224	CC	intrinsic to membrane	6	0.023	0.44
GO:0031982	CC	vesicle	29	0.037	0.44
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	29	0.036	0.44
GO:0031410	CC	cytoplasmic vesicle	29	0.037	0.44
GO:0016021	CC	integral to membrane	6	0.022	0.44
GO:0044425	CC	membrane part	6	0.082	0.84

Table 4: GO terms for genes significantly up-regulated exclusively in plants inoculated with C2 isolate compared to the uninoculated treatment. Column headings follow Table1.

GO analysis was carried out on each group of probe sets. Results for the up-regulated probe sets in plants inoculated with D1 isolate are presented in Table 5. The most significant pathways were those involved in response to stimulus and response to stress. Different pathways are present for the 676 down-regulated probe sets, such as oxygen, gas transporters and iron binding (Table 6). The pathway chitinase activity was also down-regulated in plant inoculated with D1 isolate compared to plants inoculated with the C2 isolate.

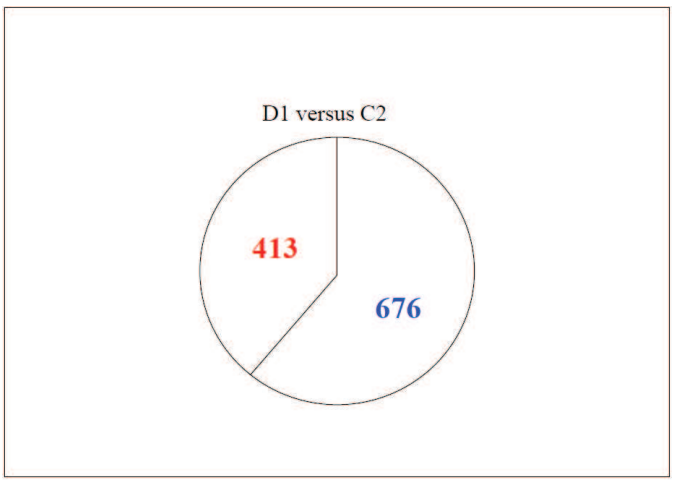


Figure 3: Blue numbers represent number of probe set that are down regulated; Red numbers represent number of probe set that are up regulated.

GO	Type	Functions	Count	p-value	FDR
GO:0050896	BP	response to stimulus	29	1.80E-05	**0.0027**
GO:0006950	BP	response to stress	21	1.90E-05	**0.0027**
GO:0006090	BP	pyruvate metabolic process	7	0.00022	**0.021**
GO:0015976	BP	carbon utilization	5	0.0024	0.17
GO:0051186	BP	cofactor metabolic process	9	0.0035	0.2
GO:0006732	BP	coenzyme metabolic process	7	0.013	0.61
GO:0042221	BP	response to chemical stimulus	11	0.017	0.61
GO:0032787	BP	monocarboxylic acid metabolic process	9	0.016	0.61
GO:0019748	BP	secondary metabolic process	6	0.02	0.63
GO:0030001	BP	metal ion transport	5	0.026	0.74
GO:0016614	MF	oxidoreductase activity	7	0.00016	**0.015**
GO:0016616	MF	oxidoreductase activity	7	8.20E-05	**0.015**

Table 5: GO terms for genes significantly up-regulated in comparison of plant roots inoculated with D1 isolate and C2 isolate. Column headings follow Table1.

Among the 1089 probe sets found to be different between the plant roots inoculated with C2 isolate and D1 isolate, we found 62 probe sets that were also differentially transcribed in the comparison between plant roots inoculated plant roots non-inoculated. These genes are listed in table 7. The GO analysis revealed on this set that genes involved in oxygen transport were the most affected.

GO	Type	Functions	Count	p-value	FDR
GO:0015671	BP	oxygen transport	5	2.20E-09	**1.50E-06**
GO:0034641	BP	cellular nitrogen compound metabolic process	25	4.70E-09	**1.60E-06**
GO:0015669	BP	gas transport	5	7.70E-09	**1.70E-06**
GO:0000096	BP	sulfur amino acid metabolic process	7	3.10E-05	**0.0052**
GO:0006350	BP	transcription	30	0.00025	**0.034**
GO:0043436	BP	oxoacid metabolic process	33	0.00083	**0.059**
GO:0006082	BP	organic acid metabolic process	33	0.00083	**0.059**
GO:0019752	BP	carboxylic acid metabolic process	33	0.00083	**0.059**
GO:0045449	BP	regulation of transcription	27	0.00088	**0.059**
GO:0009733	BP	response to auxin stimulus	8	0.00075	**0.059**
GO:0042180	BP	cellular ketone metabolic process	33	0.0011	**0.069**
GO:0032787	BP	monocarboxylic acid metabolic process	15	0.0018	**0.1**
GO:0009889	BP	regulation of biosynthetic process	28	0.0023	**0.1**
GO:0010556	BP	regulation of macromolecule biosynthetic process	28	0.0021	**0.1**
GO:0031326	BP	regulation of cellular biosynthetic process	28	0.0023	**0.1**
GO:0020037	MF	heme binding	15	1.20E-07	**3.30E-05**
GO:0046906	MF	tetrapyrrole binding	15	2.50E-07	**3.40E-05**
GO:0005506	MF	iron ion binding	21	4.70E-07	**4.30E-05**
GO:0019825	MF	oxygen binding	5	1.00E-06	**6.90E-05**
GO:0016762	MF	xyloglucan:xyloglucosyl transferase activity	5	2.00E-05	**0.0011**
GO:0030528	MF	transcription regulator activity	25	8.20E-05	**0.0038**
GO:0046872	MF	metal ion binding	55	0.00022	**0.0085**
GO:0004568	MF	chitinase activity	5	0.0004	**0.012**
GO:0016491	MF	oxidoreductase activity	30	0.00037	**0.012**
GO:0043169	MF	cation binding	55	0.00054	**0.014**
GO:0043167	MF	ion binding	55	0.00055	**0.014**
GO:0000287	MF	magnesium ion binding	6	0.0079	0.18
GO:0003700	MF	transcription factor activity	15	0.013	0.27
GO:0005488	MF	binding	139	0.015	0.29
GO:0016616	MF	oxidoreductase activity	5	0.026	0.38
GO:0004553	MF	hydrolase activity	7	0.026	0.38
GO:0046983	MF	protein dimerization activity	7	0.024	0.38
GO:0005515	MF	protein binding	34	0.022	0.38
GO:0046914	MF	transition metal ion binding	34	0.027	0.38
GO:0050662	MF	coenzyme binding	8	0.03	0.41
GO:0016614	MF	oxidoreductase activity	5	0.038	0.5
GO:0031982	CC	vesicle	127	0.0035	**0.1**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	127	0.0034	**0.1**
GO:0031410	CC	cytoplasmic vesicle	127	0.0035	**0.1**
GO:0031988	CC	membrane-bounded vesicle	127	0.0034	**0.1**
GO:0005634	CC	nucleus	39	0.0032	**0.1**
GO:0008287	CC	protein serine/threonine phosphatase complex	6	0.0018	**0.1**
GO:0031981	CC	nuclear lumen	13	0.0044	0.11
GO:0005576	CC	extracellular region	6	0.017	0.37
GO:0031974	CC	membrane-enclosed lumen	13	0.03	0.38
GO:0043233	CC	organelle lumen	13	0.028	0.38
GO:0044428	CC	nuclear part	13	0.028	0.38
GO:0005654	CC	nucleoplasm	10	0.023	0.38
GO:0070013	CC	intracellular organelle lumen	13	0.028	0.38

Table 6: GO terms for genes significantly down-regulated in comparison of plant roots inoculated with D1 isolate and C2 isolate. Column headings follow Table1.

GO	Type	Functions	Count	p-value	FDR
GO:0015671	BP	oxygen transport	5	2.50E-14	**3.70E-13**
GO:0015669	BP	gas transport	5	8.70E-14	**6.40E-13**
GO:0006810	BP	transport	7	0.0065	**0.021**
GO:0051234	BP	establishment of localization	7	0.0065	**0.021**
GO:0051179	BP	localization	7	0.007	**0.021**
GO:0019825	MF	oxygen binding	5	1.20E-11	**4.10E-10**
GO:0020037	MF	heme binding	6	4.40E-07	**6.80E-06**
GO:0046906	MF	tetrapyrrole binding	6	6.10E-07	**6.80E-06**
GO:0005506	MF	iron ion binding	6	3.00E-05	**0.00025**
GO:0043169	MF	cation binding	9	0.009	**0.043**
GO:0046872	MF	metal ion binding	9	0.0071	**0.043**
GO:0043167	MF	ion binding	9	0.009	**0.043**
GO:0046914	MF	transition metal ion binding	7	0.012	**0.048**
GO:0031988	CC	membrane-bounded vesicle	16	0.048	0.65
GO:0031982	CC	vesicle	16	0.048	0.65
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	16	0.048	0.65
GO:0031410	CC	cytoplasmic vesicle	16	0.048	0.65

Table 7: GO terms for genes commonly altered for their transcription in comparison of plant roots inoculated with D1 and C2 isolate and in comparison between mycorrhizal and non-mycorrhizal treatment. Column headings follow Table1.

Discussion

In this study, using a microarray approach, we detected which genes transcription is altered by the presence of a symbiotic partner. The direct comparison between the microarrays hybridized with the transcripts obtained from the plant roots inoculated with one or the other AMF isolates, allowed us to show that plant transcriptome can be considerably altered by genetic variability among AMF isolates.

Guimil *et al.* (2005) carried out a comparison between microarray hybridized with RNA coming from rice root inoculated with AMF or not. They found a total of 224 genes that were specifically transcribed in the presence of AMF. We observed whether transcription of these genes was altered in our experiment. We found 39 genes in common. These genes are presented in Table S7. The majority of the 39 genes showed the same trends in terms of up- and down-regulation as in the study by Guimil *et al.* (2005). Nevertheless, twelve of these were found to be regulated differently. One explanation could be the fact that we did not use the same AMF isolate as in their experiment. The AMF species *G. intraradices* shows a high level of genetic variation (Stockinger *et al.* 2010). C2 and D1 isolates, used in our study, and coming from the same field in Tänikon, Switzerland, are genetically distant from the isolate used by Guimil *et al.* (2005) (DAOM 181602). One alternative hypothesis could have been that the fungi used in the two studies may have had different effects on plant growth. Despite the fact that the plants were well colonized in both studies, no effect of the AMF on plant growth occurred in our study or in the study by Guimil *et al.* (2005) (Professor Paszkowski, personal communication). Given that the smaller genetic difference between the two isolates C2 and D1 induced considerable transcriptional differences in the plants, perhaps it is not surprising that a much more distant isolate induced transcriptional changes in many other genes to those observed in this study. Moreover, in the study of by Guimil *et al.* (2005), RNA was extracted 6 weeks after inoculation. In our experiment, we extracted the plant root RNA at 12 weeks after inoculation. This difference of 6 weeks can also contribute to the differences in the transcriptional profile. It possible that genes are not transcribed in the same way between a short time symbiosis establishment (6 weeks) and a long time (12 weeks) symbiosis establishment. Gutjahr *et al.* (2008) carried out an experiment to characterize transcription of some mycorrhizal specific genes previously identified by Guimil *et al.* (2005). These mostly

are of unknown function, except Pt11, which was already known as phosphate transporter. These genes were previously shown to be expressed only when the mycorrhizal fungus *G. intraradices* was present. In our experiment, we observed whether those genes were differentially expressed between the mycorrhizal treatment and NM treatment. We found that only three of those genes were differentially expressed between our mycorrhizal treatment and NM treatment. AM2, AM39 was found to be up-regulated and AM137 was found to be down-regulated in plant roots inoculated with D1 isolate in comparison with NM treatment (Table S8). The transcription of these genes was not found to be altered when plant roots were inoculated with C2 isolate. The study by Gutjahr *et al.* (2008) indicates that considerable changes in transcription levels of mycorrhyzal-specific genes occurs between 6 and 12 weeks that can also explain that we did not retrieved the same genes.

Carbon metabolism

During a functional AM symbiosis, the plant supplies the fungi with carbohydrates. Two studies propose hexoses as the molecule released by the plant to the fungus (Shachar-Hill *et al.* 1995; Bago *et al.* 2003). Lipids are also propose as molecule released by the plant to the fungus (Bago *et al.* 2003)

We observed that plant sugar metabolism is altered and that many genes involved in sugar metabolism are differentially expressed. The GO analysis confirmed that several pathways involved in sugar metabolism were up-regulated in mycorrhizal plants compared to non-mycorrhizal plants. Moreover, all these processes belong to related pathways (Figure S2). The mycorrhizal specific genes found in Guimil *et al.* (2005) and characterized in Gutjahr *et al.* (2008) do not comprise this group of genes involved in sugar metabolism and therefore should be considered as sensitive to the AM symbiosis.

Response of the host

A fundamental study proposed differential transcription of plant chitinase gene families according to whether the plant is colonized by a mutualistic symbiont or a pathogen (Salzer *et al.* 2000). Within this frame, transcription levels of chitinase families I, II and IV are enhanced

by bacteria and also by pathogenic fungi. However, mycorrhizal fungal symbionts do not cause up regulation in transcripts of chitinase encoding genes in classes I, II and IV. Conversely, chitinase class III enzymes are proposed to be a mycorrhizal specific group in *Medicago truncatula* (Salzer *et al.* 2000).

In our study, four probe sets corresponding to class I chitinases; Chia4a, Chit 7, Chit 14, and Os.1721.S1_a_at were commonly down-regulated in both mycorrhizal treatments and the pathway involving these enzymes is down-regulated. No transcription of genes encoding chitinase class III enzyme was detected in our study.

Moreover, other pathways involving hydrolase enzymes were down-regulated, showing that plant seems to recognize the fungus as a symbiont. These results are consistent with recent finding by Kloppholz *et al.* (2011) where an AMF effector protein has been found that interacts with a plant protein stopping the activation of some plant defence responses. Finally, many GO terms involved in response to stress, other organisms or defence were down-regulated (Figure S3).

Our study is consistent with other studies in that pathways involved in chitinases or hydrolases are altered between the mycorrhizal and the non-mycorrhizal treatment. However, results also revealed that pathways involved in responses to stimuli or stress are down-regulated in roots inoculated with the D1 isolate compared to those inoculated with isolate C2 (Figures S3, S4, S5, S6). Genes belonging to these GO pathways are involved in different functions such as responses to iron, water privation or desiccation. These results showed that the genetic variability among AMF isolate could induce a different transcriptomic response in the plant, and therefore reveals that plant gene transcription might be specifically altered by given isolates.

Inorganic Phosphate

Interestingly, we found no difference in gene transcription in our study for genes encoding phosphate transporters. Even transcription of the gene encoding phosphate transporter 11 (Pt11), which is specific to the AM symbiosis, was not altered in our study. We found that a glucose-6-phosphate/phosphate translocator gene is up-regulated in comparison between plant

roots inoculated with C2 isolate and with the NM treatment. However, this gene is not known to be mycorrhizal specific. The absence of differential phosphate transporter transcription in our data set may be because no difference in plant growth was detected at 12 weeks after inoculation. Guimil *et al.* (2005) found that Pt11 transcription was increased although they extracted RNA 6 weeks after inoculation.

Conclusion

In this study, we investigated the effect of AMF on rice transcriptome, using an alternative approach. In addition to the traditional comparison between mycorrhizal treatments versus NM treatment, we added a comparison between roots inoculated with two genetically different AMF isolates. This approach allowed us to reveal that AMF from the same species, even from the same population, can act differently on the transcriptome of their host and this shows the complexity of the AM symbiosis at the molecular level.

Acknowledgments

This work was supported by grants 31000AO-105790/1 and 31003A-127371 from the Swiss National Science Foundation. We thank U. Paszkowski for providing rice seeds, J. Bonvin for assistance with fungal inoculation and the rest of our group for their helpful comments on the manuscript. We also thank the Genomic Technology Facility of the University of Lausanne for their assistance for the microarray hybridization and the quantitative PCR.

Supplementary files

Microarray validation

We used several different approaches to check the validity of the transcription profiles generated from the microarrays. These were:

1) Quantitative PCR

Quantitative PCR was carried out on selected genes. New RNA extractions were performed on the original samples. We used the RNeasy Plant mini kit (Qiagen Inc®) to extract RNA following the manufacturer's recommendations. Precipitation with LiCl was also performed (Ambion®)

Quantitative PCRs were performed on three genes that were detected in the microarrays because of their fold change expression between the isolates and the NM treatment: Os.52079.1.S1 (unknown function, but contain a domain mainly found in plants), OS.7947.1.S1 (Lichenase-2 precursor) and Os.5319.1.S1 (Phosphoglucomutase, and play a role in carbohydrate metabolism). Primer 3 software was used to design the primers. PCR conditions and mixs were adopted from Gutjahr *et al*. (2008).

Power Sybr® green was used as the label (Applied Biosystems) for quantitative PCR and each sample was run in triplicate. Quantitative PCR was performed on Prism AB7900 quantitative PCR machine. Ubiquitin transcript levels were used to normalize the data. Transcription values were calculated with SDS 2.3 software (Applied Biosystems) and qBASE Plus software (Hellemans *et al.* 2007).

Quantitative PCR confirmed the microarrays results (Figure S1, Table S1). For example, Os.52079.1.S1 was down-regulated by a fold change of 11 and 10 for C2 and D1 respectively on microarray. Quantitative PCR showed a similar fold change and differences were statistically different. Quantitative PCR results on Os.5319.1.S1 transcript level were congruent with the microarray results although the differences were not significant.

Analyses on gene transcription levels were performed for each gene separately. When ANOVAs were significant, we carried-out a multiple comparison test using Tukey-kramer HSD test with an alpha level of 0.05.

All analyses were performed with the statistical program JMP® version 5.0 (SAS Institute Inc., Cary, NC, USA).

2) Gene Ontology analysis

We made gene Ontology (GO) analysis on the microarray results with the software AgriGO (website: http://bioinfo.cau.edu.cn/agrigo/). The GO analyses were performed on the probe sets with fold changes 2 or more, or -2 or less. The GO terms generated in the GO analysis tests whether a significant number of genes that are up- or down-regulated group together into a given metabolic pathways, cellular functions or the cell cycle. If the microarray data were not reliable then significant GO terms should not be generated in the analysis. The significance of the GO terms is determined by a false discovery rate (FDR) statistic. Many of the significant GO terms in our analysis also contained a large number of significantly up- or down-regulated probe sets, which would not be expected if the microarray data were simply showing random patterns. The GO analysis also reveals positive and negative regulation among significant GO terms. Again significant positive and negative relationship between GO terms would not be expected.

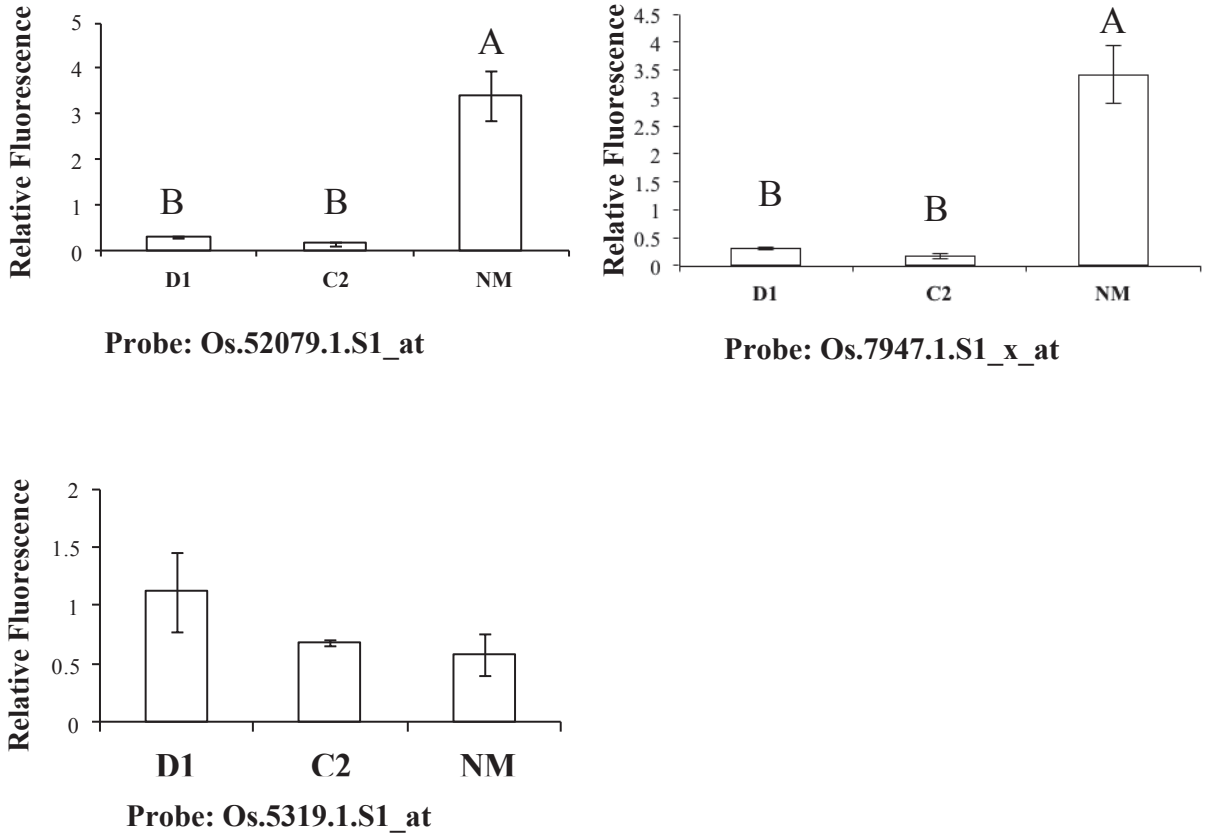


Figure S1: Quantitative PCR for the validation of microarray using three probe sets.

Error bars ± 1 SE. Bars topped by different letters are significantly different at $p \leq 0.05$.

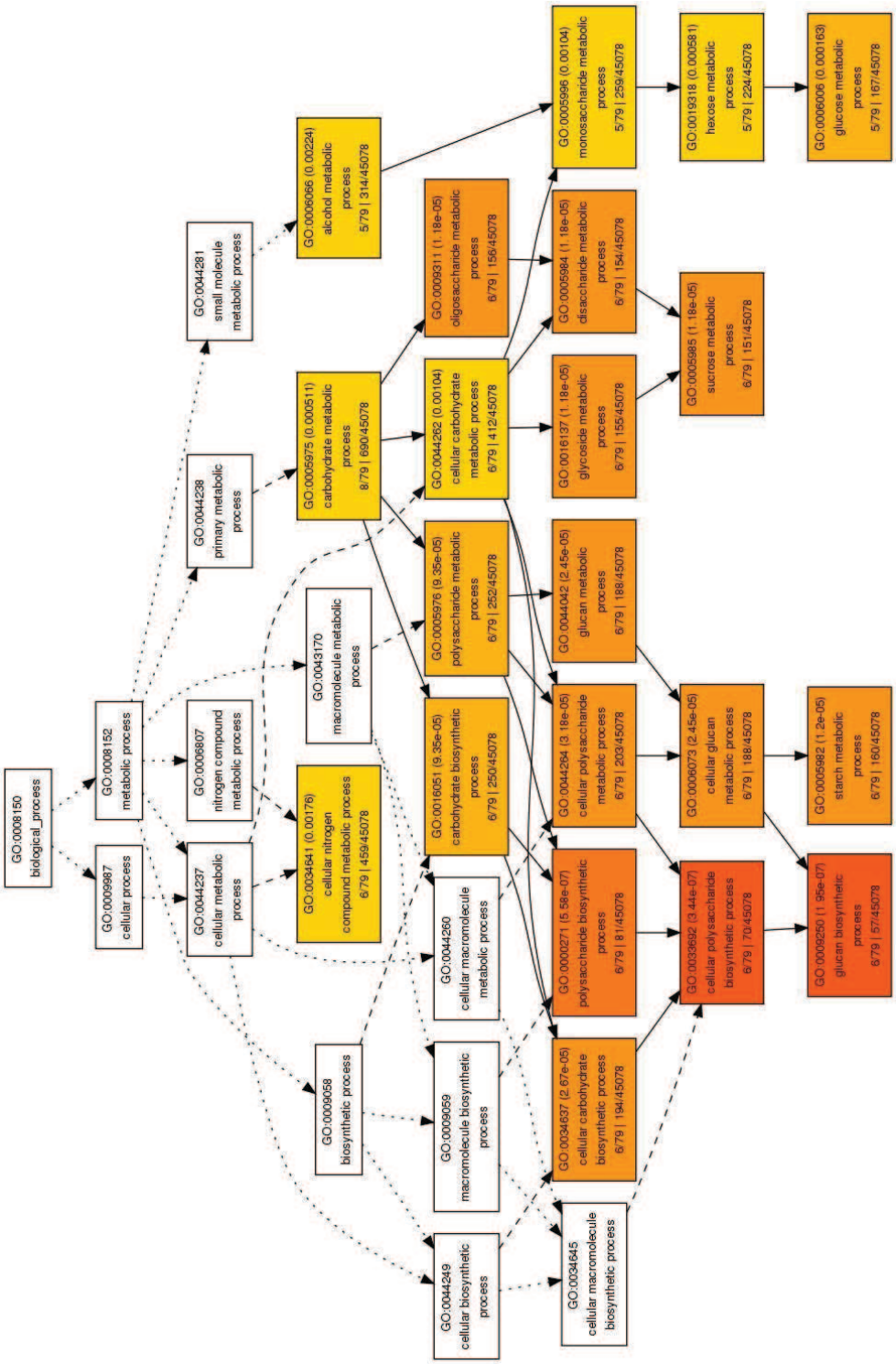


Figure S2: Graphical representation of the significant GO terms representing biological processes in the GO analysis for the probe sets up-regulated in the comparison between inoculated and non-inoculated plants.

Each box represents a GO term. FDRs are associated with each GO term. The more the FDR is significant, the redder the box. White boxes mean that no gene is found in this GO term. Dotted lines between two GO terms mean that a negative regulation occurred between the two GO terms. Continuous lines between tow GO terms mean a positive regulation between the two GO terms.

Figure S3: Graphical representation of the significant GO terms representing molecular functions in the GO analysis for the probe sets up-regulated in the comparison between plants inoculated with isolate C2 and plants inoculated with isolate D1.

Each box represents a GO term. FDRs are associated with each GO term. The more the FDR is significant, the redder the box. White boxes mean that no gene is found in this GO term. Dotted lines between two GO terms mean that a negative regulation occurred between the two GO terms. Continuous lines between tow GO terms mean a positive regulation between the two GO terms.

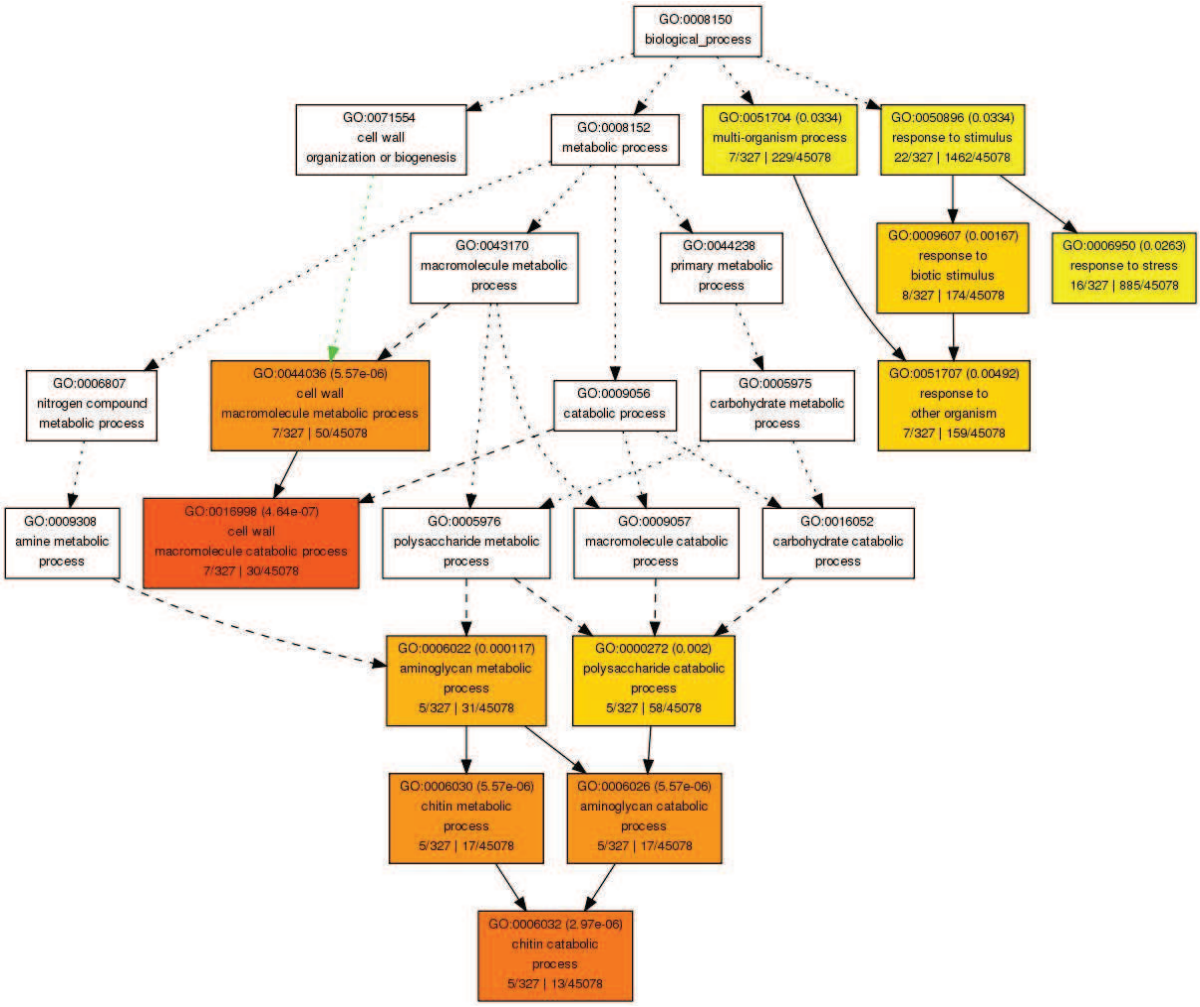


Figure S4: Graphical representation of the significant GO terms representing biological processes in the GO analysis for the probe sets down-regulated in the comparison between plants inoculated with isolate C2 and plants inoculated with isolate D1.

Each box represent a GO terms. FDR are associated in each GO term. More the FDR are significant, and more the boxes are red. White boxes mean that no gene is found in this GO term. Dotted lines mean that a negative regulation occurred between the two GO terms. Completed lines mean a positive regulation between the two GO terms.

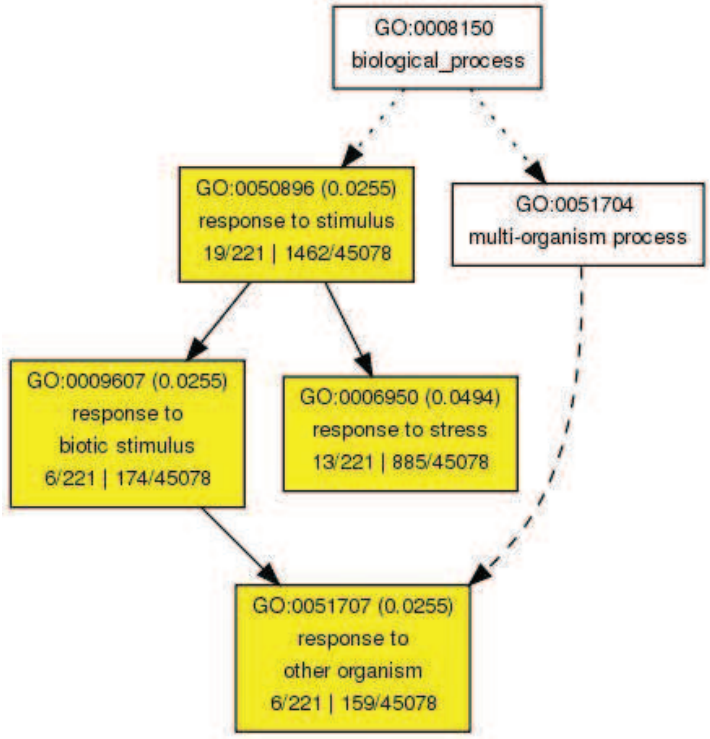


Figure S5: Graphical representation of the significant GO terms representing biological processes in the GO analysis for the probe sets down-regulated in the comparison between plants inoculated with isolate C2 and plants inoculated with isolate D1.

Each box represent a GO terms. FDR are associated in each GO term. More the FDR are significant, and more the boxes are red. White boxes mean that no gene is found in this GO term. Dotted lines mean that a negative regulation occurred between the two GO terms. Completed lines mean a positive regulation between the two GO terms.

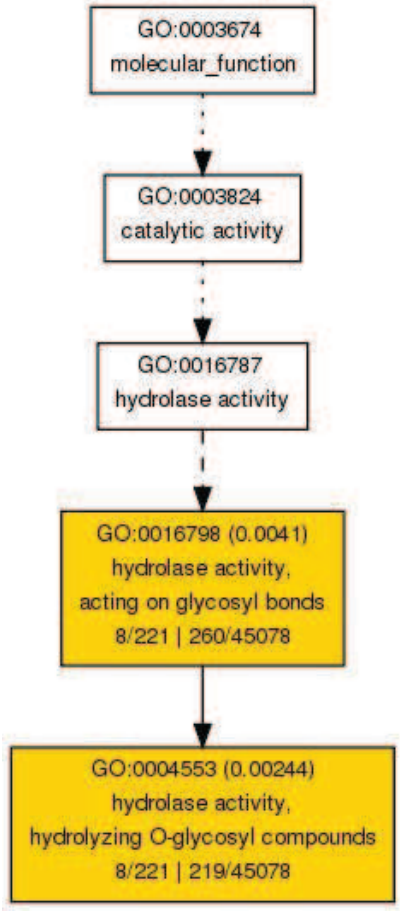


Figure S6: Graphical representation of the significant GO terms representing molecular functions in the GO analysis for the probe sets down-regulated in the comparison between plants inoculated with isolate C2 and plants inoculated with isolate D1.

Each box represent a GO terms. FDR are associated in each GO term. More the FDR are significant, and more the boxes are red. White boxes mean that no gene is found in this GO term. Dotted lines mean that a negative regulation occurred between the two GO terms. Completed lines mean a positive regulation between the two GO terms.

Source	d.f.	SS	F	*p*	Same trends as in microarray data
Os.52079.1.S1_at	**5**	**16.034**	**7.403**	**0.0022***	**Yes**
Os.7947.1.S1_x_at	**5**	**12.293**	**59.813**	**<0.001***	**Yes**
Os.5319.1.S1_at	**5**	**2.126**	**2.323**	**0.1075**	**Yes**

Table S1: Analysis of variance on plant transcription of three genes for the microarray validation.

Probe sets	Functions
Os.16044.1.S1_at	Similar to MtN3 protein precursor
Os.51775.1.S1_at	Conserved hypothetical protein
OsAffx.11871.1.S1_s_at	EF-Hand type domain containing protein

Table S2: Genes differentially transcribed between Mycorrhizal treatment and NM treatment by a fold change of two. These genes are up-regulated for the plant inoculated with isolate D1, but down-regulated for those inoculated with isolate C2.

Probe sets	Functions
Os.54146.1.S1_at	Nitrate reductase, putative, expressed
Os.49093.1.S1_at	Similar to BCH2
OsAffx.22999.1.S1_at	Hypothetical protein
Os.7879.2.S1_at	Glutamine synthetase, chloroplast precursor, putative, expressed
Os.7317.2.S1_at	Protease inhibitor/seed storage/LTP family protein, expressed
Os.4863.1.S1_at	Siroheme synthase, putative, expressed
Os.12498.2.S1_at	Expressed protein
Os.20614.1.S1_at	F-box protein interaction domain containing protein, expressed
Os.49826.1.S1_at	Expressed protein
OsAffx.27299.1.S1_at	Non-symbiotic hemoglobin 2, putative
Os.8823.1.S1_at	Expressed protein
Os.43929.1.S1_s_at	Expressed protein
Os.10266.1.S1_at	Expressed protein
Os.16037.1.S1_at	Unknown
Os.170.3.S1_at	Ferredoxin-6, chloroplast precursor, putative, expressed

Table S3: Probe sets differentially transcribed between the mycorrhizal treatment and NM treatment by a fold change of two. These probe sets are down-regulated in the plant inoculated with isolate D1, but up-regulated for those inoculated with isolate C2.

Probe sets involved in "transporter activity"	Function
Os.47946.1.S1_s_at	Glucose-6-phosphate/phosphate translocator 2, chloroplast precursor, putative, expressed
Os.50896.1.S1_at	Amino acid carrier, putative, expressed
Os.27228.1.S1_at	Hexose carrier protein HEX6, putative, expressed
Os.35495.1.S1_s_at	ADP-ribosylation factor, putative, expressed
OsAffx.24678.2.S1_at	Aquaporin PIP2.2, putative, expressed
Os.10155.1.S1_at	Amino acid selective channel protein, putative, expressed
Os.35495.1.S1_at	ADP-ribosylation factor, putative, expressed
OsAffx.19332.1.S1_at	Solute carrier family 2, facilitated glucose transporter member 8, putative, expressed
Os.8916.1.S1_at	Permease I, putative, expressed

Table S4: Nine probe sets involved in the GO terms Transport (GO:0006810), with their putative function. These significant probe set are exclusively up-regulated in plants inoculated with isolate C2 when compare to the NM treatment.

GO	Type	Functions	Count	p-value	FDR
GO:0009579	CC	thylakoid	12	4.60E-06	**0.00076**
GO:0034357	CC	photosynthetic membrane	8	0.00045	**0.038**
GO:0016020	CC	membrane	66	0.001	**0.058**
GO:0009507	CC	chloroplast	17	0.0018	**0.075**
GO:0044435	CC	plastid part	9	0.0024	**0.081**
GO:0044434	CC	chloroplast part	8	0.0052	0.14
GO:0031090	CC	organelle membrane	11	0.0064	0.15
GO:0055035	CC	plastid thylakoid membrane	5	0.011	0.2
GO:0009535	CC	chloroplast thylakoid membrane	5	0.01	0.2
GO:0042651	CC	thylakoid membrane	5	0.015	0.26
GO:0009534	CC	chloroplast thylakoid	5	0.025	0.37
GO:0031976	CC	plastid thylakoid	5	0.028	0.38
GO:0031984	CC	organelle subcompartment	5	0.03	0.38
GO:0044436	CC	thylakoid part	5	0.036	0.43

Table S5: GO terms for genes significantly up-regulated exclusively in plants inoculated with isolate D1 compared to the uninoculated treatment. Column headings follow Table1.

GO	Type	Functions	Count	p-value	FDR
GO:0034641	BP	cellular nitrogen compound metabolic process	26	1.80E-07	**0.00013**
GO:0009066	BP	aspartate family amino acid metabolic process	10	0.00031	0.11
GO:0006118	BP	electron transport	15	0.0011	0.26
GO:0006350	BP	transcription	33	0.0021	0.38
GO:0010556	BP	regulation of macromolecule biosynthetic process	33	0.0042	0.44
GO:0031326	BP	regulation of cellular biosynthetic process	33	0.0046	0.44
GO:0009889	BP	regulation of biosynthetic process	33	0.0046	0.44
GO:0045449	BP	regulation of transcription	30	0.0048	0.44
GO:0000096	BP	sulfur amino acid metabolic process	5	0.0061	0.48
GO:0015833	BP	peptide transport	7	0.0079	0.48
GO:0010468	BP	regulation of gene expression	33	0.007	0.48
GO:0006857	BP	oligopeptide transport	7	0.0079	0.48
GO:0019219	BP	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	31	0.01	0.57
GO:0051171	BP	regulation of nitrogen compound metabolic process	31	0.011	0.57
GO:0019318	BP	hexose metabolic process	9	0.015	0.74
GO:0080090	BP	regulation of primary metabolic process	33	0.019	0.78
GO:0051179	BP	localization	40	0.018	0.78
GO:0009060	BP	aerobic respiration	5	0.019	0.78
GO:0044271	BP	cellular nitrogen compound biosynthetic process	11	0.022	0.83
GO:0006804	BP	peroxidase reaction	5	0.023	0.84
GO:0031323	BP	regulation of cellular metabolic process	33	0.024	0.84
GO:0060255	BP	regulation of macromolecule metabolic process	33	0.026	0.86
GO:0019222	BP	regulation of metabolic process	33	0.042	1
GO:0030528	MF	transcription regulator activity	32	1.30E-05	**0.0039**
GO:0016762	MF	xyloglucan:xyloglucosyl transferase activity	5	6.90E-05	**0.0069**
GO:0046872	MF	metal ion binding	70	6.90E-05	**0.0069**
GO:0016491	MF	oxidoreductase activity	38	0.00012	**0.009**
GO:0043169	MF	cation binding	70	0.00021	**0.0092**
GO:0043167	MF	ion binding	70	0.00022	**0.0092**
GO:0005506	MF	iron ion binding	19	0.00021	**0.0092**
GO:0003700	MF	transcription factor activity	23	0.00031	**0.012**
GO:0046906	MF	tetrapyrrole binding	11	0.0014	**0.046**
GO:0016616	MF	oxidoreductase activity,	8	0.0016	**0.047**
GO:0016614	MF	oxidoreductase activity, acting on CH-OH group of donors	8	0.003	**0.074**
GO:0020037	MF	heme binding	10	0.0029	**0.074**
GO:0000287	MF	magnesium ion binding	7	0.0076	0.18
GO:0005509	MF	calcium ion binding	11	0.009	0.19
GO:0046914	MF	transition metal ion binding	45	0.01	0.2
GO:0003677	MF	DNA binding	36	0.018	0.33
GO:0016769	MF	transferase activity, transferring nitrogenous groups	5	0.03	0.53
GO:0016684	MF	oxidoreductase activity, acting on peroxide as acceptor	5	0.036	0.54
GO:0004601	MF	peroxidase activity	5	0.036	0.54
GO:0005488	MF	binding	172	0.035	0.54
GO:0004252	MF	serine-type endopeptidase activity	6	0.044	0.62

Table S6: GO terms for genes significantly down-regulated exclusively in plants inoculated with isolate D1 compared to the uninoculated treatment. Column headings follow Table1.

Genes	Functions	K vs NM DN	K vs NM UP	F vs NM DN	F vs NM UP
OsAM2	Hypothetical protein		x		
OsAM4	Putative germin-like protein		x		
OsAM5	Hypothetical protein		x		
OsAM6	Putative epoxide hydrolases		x		
OsAM7	Putative disease resistance protein RPM1				
OsAM9	Putative germin-like protein		x		
OsAM16	Cysteine protease		x		
OsAM19	Putative serine carboxypeptidase		x		
OsAM24	Putative cytochrome P450		x		
OsAM38	Putative 1-deoxyxylulose 5-phosphate synthase			x	
OsAM39	Cysteine protease		x		
OsAM43	Putative plasma membrane H+-ATPase		x		
OsAM47	Putative beta-N-acetylglucosaminidase		x		
OsAM49	Hypothetical protein		x		
OsAM67	Hypothetical protein		x		
OsAM69	Diphosphonucleotide phosphatase-like protein				
OsAM77	AP2 domain-containing transcription factor-like	x			
OsAM84	Putative multidrug resistance protein 1 homolog (MDR-like ABC transporter)		x		
OsAM88	FH protein interacting protein FIP1-like		x		
OsAM89	Putative class IV chitinase (CHIV)			x	
OsAM90	Ethylene-responsive element binding factor 3	x			
OsAM101	Putative bifunctional nuclease	x			
OsAM106	Hypothetical protein		x		
OsAM119	Putative enolase				
OsAM122	Putative P450 monooxygenase		x		
OsAM134	Putative cysteine synthase		x		
OsAM136	Hypothetical protein		x		x
OsAM137	Zinc finger transcription factor ZF1	x			
OsAM146	Phytoene synthase				
OsAM163	Putative cinnamoyl CoA reductase		x		
OsAM175	Hypothetical protein			x	
OsAM176	Putative delta 12 oleic acid desaturase	x		x	
OsAM178	Hypothetical protein		x		
OsAM187	Ent-CDP synthase OsCyc2	x		x	
OsAM187	Ent-CDP synthase OsCyc2				
OsAM189	Putative flavonoid 3',5'-hydroxylase		x		
OsAM195	Putative delta 12 oleic acid desaturase			x	
OsAM198	Hypothetical protein	x		x	
OsAM204	Hypothetical protein	x			
OsAM213	Hypothetical protein		x		
OsAM215	Putative diphosphate-fructose-6-phosphate 1-phosphotransferase alpha chain				
OsAM216	Hypothetical protein	x		x	
OsAM218	Hypothetical protein	x			
OsAM219	Hypothetical protein	x			
OsAM223	Putative chitinase			x	

Table S7: Gene commonly altered in their transcription in our experiment and in Guimil *et al.* experiment. Gene in bold are down-regulated in Gumil *et al.*(2005). NM: non-Mycorrhizal; DN: Down-regulated; UP: Up-regulated

Genes	Putative function	Regulation
AM2	hypothetical protein (O. s.), proteinase inhibitor I13 in potato 1	Up-regulated
AM39	cysteine peptidase, protease family C1	Up-regulated
OsAM137	Zinc finger transcription factor ZF1	Down-regulated

Table S8: Comparison of our results with the results of Gutjahr *et al.*, which characterized 38 mycorrhizal specific genes. Only three of these genes were found in our analysis when we compared the genes found by comparison of mycorrhizal treatment versus non-mycorrhizal treatment.

Chapter 3

Segregation in arbuscular mycorrhizal fungi alters parts of the rice transcriptome

Alexandre Colard, Caroline Angelard and Ian R. Sanders

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

Abstract

Arbuscular mycorrhizal fungi (AMF) are important symbionts of most plant species, promoting plant diversity and productivity. They play a role in plant nutrients acquisition, especially phosphate. In exchange, AMF obtain photosynthetically fixed carbon. Evidence exists suggesting that AMF contain populations of genetically different nucleotypes coexisting in a common cytoplasm. Segregation, which is the random distribution of nuclei at spore formation, has recently been demonstrated in AMF *Glomus intraradices*. Studies have shown that segregation in AMF could alter the transcription of symbiosis specific genes. Because others studies have shown that different AMF individual could have different effects on plant genes transcription, we investigated the effect of segregation in AMF on plant genes transcription. We found that pathways involved in plant defences and in carbohydrates are the most altered. We also showed that some pathways were specific depending of the inoculated segregated lines. Moreover, we detected that the transcription of Pt11, an AM specific gene, was altered in plants inoculated with two segregated lines. In overall, results showed that segregation in AMF can affect plants transcriptome.

Introduction

Arbuscular mycorrhizal fungi (AMF) form symbioses with more than 80% of land plants. The mycorrhizal symbiosis can promote plant diversity and productivity (van der Heijden *et al.* 1998). AMF supply nutrients to plants, mainly phosphate (Smith *et al.* 2008) but also nitrogen (Hodge *et al.* 2001) and other ions, such as copper, zinc or iron (Harrison 1999; Smith *et al.* 2008). In exchange, plants provide carbon to the fungi (Jakobsen *et al.* 1990). AMF can also increase resistance to various environmental stresses such as herbivory (Gange *et al.* 2003), heavy metals (Ricken *et al.* 1996) or drought (Ebel *et al.* 1996).

Some AMF have been shown to be heterokaryotic, containing populations of genetically different nuclei within one fungus (Kuhn *et al.* 2001; Hijri *et al.* 2005). Segregation of genetically different nuclei was shown to occur in AMF (Angelard *et al.* 2010). Segregation in AMF results in different distributions of nuclei among newly created spores during spore formation (Angelard *et al.* 2010). This results in new spores being qualitatively different, because not all nuclei will necessarily be inherited in all new spores, and quantitatively, because the nuclei may not have the same frequency in all newly formed spores (Angelard *et al.* 2010). They also found that this process can alter the transcription of symbiosis specific genes. Moreover, Segregation in AMF also affects how the fungus colonizes the host (Angelard *et al.* 2011).

Given that segregation in AMF can lead to transcriptional changes in genes known to be involved in the symbiosis, a larger scale study on the effect of segregation on genome-wide plant gene transcription is warranted. To date, several studies have investigated the effects of AMF on different plants transcriptome (Liu *et al.* 2003; Wulf *et al.* 2003; Guimil *et al.* 2005; Hohnjec *et al.* 2005; Fiorilli *et al.* 2009; Grunwald *et al.* 2009; Guether *et al.* 2009; Breuillin *et al.* 2010). In 2005, a study compared the transcriptome profiles of rice when plants were inoculated with AMF or were left uninoculated (Guimil *et al.* 2005). This study revealed a set of genes that are considered as mycorrhizal specific as well as many others that are up and down-regulated in the symbiosis. In chapter 2, we showed that genetically different AMF isolates coming from a same population could alter many different metabolic pathways in rice in addition to pathways involved in plant growth. This study has demonstrated that closely related AMF isolates can act differently on the transcriptome of their host but also has shown

the complexity of the AM symbiosis at the molecular level. Study of the effects of AMF segregated lines on their host is crucial for the understanding of the AM symbiosis.

In view of these results, we investigated the effects of segregation in AMF on the genome-wide rice transcriptome profiles. Using affimetrix microarrays, we hybridized total RNA coming from rice roots with either one of three parental lines or 12 segregated lines that were progeny of the parental lines. We showed that segregation in AMF affects the transcription of many rice genes.

Material and methods

Origin of fungal lines

Fifteen *Glomus intraradices* lines were used in this study. In a previous study, two pairs of AMF lines (C2xC3 and C3xD1) were co-cultured together to obtain crossed lines exhibiting bi-parental inheritance (Croll *et al.* 2009). The AMF lines C2, C3 and D1 originate from one agricultural field in Tänikon, Switzerland (Koch *et al.* 2004). In the present study, two crossed lines (S3 and S4) coming from the pairing C2xC3 and one crossed line (Sc2) coming from the pairing C3xD1 were used. Sc2 in the present study corresponds to Sc in the study of Croll *et al.* (2009). These isolates (Sc2, S3 and S4) are subsequently referred to as parental lines. Twelve segregated lines, Sc2a-f, S3a-c and S4a-c, were lines generated taking single spores from the parental lines Sc2, S3 and S4, respectively and establishing new cultures from each spore. They were shown to be qualitatively or quantitatively genetically different from each other and from the parental lines (Angelard *et al.* 2010). Some of them were also shown to differ in their phenotypic traits and differentially affected plant growth (Angelard *et al.* 2010; Angelard *et al.* 2011). These lines were cultivated in standard conditions with Ri T-DNA transformed carrot roots (Becard *et al.* 1988) and maintained through clonal subculturing with the same protocol as in Koch *et al*. (2004). We used *Oryza sativa* (Nipponbare) as plant host for these experiments because this variety had previously been used for studies of AM induced transcriptional changes.

Plant inoculation and harvesting

The substrate for growing plants was composed of 60% sand and 40% loam and was autoclaved twice at 120°C and left open to the air for two weeks in order to oxidize any toxic compounds. The seeds of rice cv. Nipponbare (kindly supplied by Uta Patzowski's group, University of Lausanne) were surface-sterilized in Sodium hypochlorite NaClO 2% for 10 minutes, washed repeatedly in sterile water and placed on filter paper for germination. Eight days after germination, seedlings were put on trays with moist vermiculite, and were watered for another five days. Each pot was filled with 450 ml of soil and one rice seedling was planted in each pot. AMF spores were extracted from four month old cultures using the same method as in Koch *et al.* (2006). The plant inoculation protocol was the same as that performed by Angelard *et al.* (2010). Each plant was inoculated with 500 spores of one of the AMF lines. Spores were suspended in water in a tube. The tube was shaken and then number

of spores per ml was calculated under the binocular microscope. This allowed us to calculate how many ml of suspension would contain 500 spores. Some plants were not inoculated and served as the non-mycorrhizal control. The three parental lines (S3, Sc2 and S4) and the 12 segregated lines (S3a, S3b, S3c, Sc2a, Sc2b, Sc2c, Sc2d, Sc2e, Sc2f, S4a, S4b, and S4c) were used, plus an uninoculated plant treatment. Ten replicate plants were established per treatment, giving a total of 150 pots plus 10 pots for the non-mycorrhizal treatment. The position of the pots in the greenhouse was randomized. Each pot was watered with 50-75 ml of tap water every two days. Plants were fertilized twice during the experiment with a full strength Hoaglands solution containing no phosphorous. Potassium concentration was adjusted by adding KCl. Growth conditions in the greenhouse were 16 hours daylight, and temperatures vary between 18°C and 30°C. Under these conditions, *O. sativa* of the Nipponbare variety are expected to go through the following developmental stages: germination, seedling establishment, tillering, stem elongation, booting or panicle initiation and heading (Yamamoto *et al.* 2000).

After 12 weeks, the shoots of each plant were harvested, dried at 80°C for two days and weighed. The roots were immediately washed and separated in two parts. Two thirds of the roots were frozen in liquid nitrogen for subsequent RNA extraction and stored at -80°C. The remaining one third was stored in 50% ethanol and then stained with trypan blue as described in Munkvold *et al.* (2004) in order to check for fungal colonization. All the non-mycorrhizal (NM) plants were found to be AMF free. Nine plants died during the experiment and no colonization by AMF could be detected in six plants that had been inoculated with AMF. These fifteen plants were randomly distributed among all treatments and were removed of the experiment. This left 145 individuals that were used for further analyses.

RNA extraction

RNA extractions were performed directly on frozen roots. Roots were ground with a pestle in liquid nitrogen. The RNeasy Plant mini kit (Qiagen Inc®) was used to extract RNA following the manufacturers recommendations. Precipitation with LiCl was also performed (Ambion®) to remove sugars present in plant roots. DNase treatment was performed with DNase I grad (Invitrogen®) following the manufacturer's recommendations. A PCR on extracted RNA was performed to amplify the housekeeping genes cyclophilin and ubiquitin to be sure that there was no DNA contamination, with the following conditions: 94°C for 5 minutes, 35 cycles [94°C 20 seconds, 58°C 20 seconds, 72°C 50 seconds] 72°C for 7 minutes.

Microarray hybridization

RNA amplification and GeneChip rice genome array hybridization (also known as affimetrix 57) were carried out at the microarray facility of the Center for Integrative Genomics at the University of Lausanne according to the following protocol. Because a large quantity of RNA was needed, WT-Ovation™ Pico RNA Amplification System from NuGEN™ was used following the manufacturer's recommendations to amplify plant RNA and transcript it into cDNA. Labelling of the cDNA was done with the FL-Ovation™ cDNA Biotin Module V2 from NuGEN™ following the manufacturer's recommendations. At this step, three samples of each treatment were merged to form three pooled replicates per treatment. One chip was used to hybridize with each of the pooled replicate samples in each treatment, making a total of 48 chips in the complete experiment. Affymetrix Rice Genome arrays (Affymetrix, Santa Clara , CA, USA) were hybridized with 5μg of biotinylated target, at 45°C for 17 hours washed and stained according to the protocol described in Affymetrix GeneChip® Expression Analysis Manual (Fluidics protocol FS450_0004). The arrays were scanned using the GeneChip® Scanner 3000 7G (Affymetrix).

Quality control of microarray hybridization

After image acquisition, the scripts Affy (Gautier *et al.* 2004), AffyPLM (Irizarry *et al.* 2003) and Limma (Smyth 2005) implemented in bioconductor software (Gentleman *et al.* 2004), R language (Team 2008), were used to extract the information from the raw data files. No potential artifacts or background gradients were detected. Since no technical problems were found, data from all the chips used for hybridization in our experiment were further analyzed.
Using customized scripts developed for this project (available upon request) boxplots for unnormalized data were printed as a first step in the hybridization quality control. Additionally, the normalized unscaled standard error (NUSE) procedure was used as a quality control measure. Raw and background intensity images (probe level model residuals (PLM)) of the chips were printed to check for quality and spatial heterogeneity inside and between the chips. Normalization was applied via the Robust Multi-Array method (RMA) which uses a convolution background correction, quantile normalization and summarization (median polish) using a robust multi-array linear model. The efficiency of the normalization procedure, revealed in the boxplots, allows further comparisons using linear models and t-

tests, which are based on a normal distribution of the data. After normalization, expression files were extracted and printed to be used as input data for further analysis.

Gene Ontology Analysis

Probe sets were selected based on their fold change (FC). Probe sets were considered to be affected by treatment when FC was greater than two. Usual approaches use the Benjamini-Hochberg False Discovery Rate (FDR) statistic to detect affected probe sets. Because the parental isolates were obtained by crossing of isolates originate from the same population, and that segregated lines are also very closely related with parental lines, almost no FDR values coming from the basic analysis from Limma package was significant. Therefore, we based level changes on FC as a signal of presence/absence. We selected the threshold of 2 because the low number of probe sets with threshold greater than 3 or 4 did not permit us to perform the following analysis.

The lists of probe sets affected by the treatments were used to carry out a singular enrichment analysis (SEA) in AgriGO, a web-based tool and database to perform Gene Ontology (GO) analysis, especially for agriculturally relevant species (Du *et al.* 2010). GO is used widely in functional annotation and enrichment analysis. The enrichment analysis classifies probe sets into different GO terms, and give as an output any pathways that are significantly over represented in the microarray data. GO project is a major bioinformatics initiative with the aim of standardizing the representation of genes and gene product attributes across many different species and databases. The well-annotated Gramene database (rice gramene locus) was used as a background for our lists corresponding to *Oryza sativa* (var. japonica). At the same time, a fisher statistical test method with a Benjamini-Hochberg FDR correction was implemented. The significance level was set up at 0.05 and the minimum number of mapping entries was five. The ability to retrieve enriched pathways with the different groups of up or down-regulated probe sets is an indication of the reliability of the microarray data output.

Microarray validation

We selected probe sets that were up- or down-regulated, and we carried out quantitative PCR on each of them. Quantitative PCR results were congruent with microarray results (see supplementary file, Figure S1 and Table S1).

Results

Effect of parental isolates and segregated lines on plant growth

Analysis of variance was performed on the shoot dry weight. There was no significant difference observed between inoculated plants and uninoculated plants or among plants inoculated with different AMF lines (data not shown). Colonization was checked in each plant but not measured. All plants used for the analysis were well colonized.

Fifteen independent treatments were performed in triplicate and hybridized to microarrays to compare gene expression of plant roots inoculated with AMF parental lines, Sc2, S3 and S4 with their respective AMF segregated lines S3a-c, Sc2a-f, S4a-c. Moreover, one treatment without AMF inoculation was performed as negative control and hybridized to microarrays.

Overall effects of the inoculation treatments on rice gene transcription

As described in the material and methods, we used probe sets with FC>2 values as a threshold. We identified 47 probe sets that were commonly and differentially transcribed between plants inoculated with AMF and plants that were uninoculated. Among these 47 probe sets, 24 were down regulated, 23 were up-regulated in inoculated plants. GO analysis on these probe sets did not permit us to identify enrichment in specific metabolic pathways. Therefore, we decided to analysis all the 47 probe sets together. GO analysis results revealed that pathways involved in carbohydrate and polysaccharide metabolic processes were altered as well as pathways involved in the responses to stress and responses to stimuli (Table 1). Two probe sets were up-regulated in plants inoculated with the parental isolate S4 and segregated lines S4a, S4b and S4c compared to uninoculated plants, whereas these two probe sets were down-regulated in the other plants inoculated with the other AMF lines. These two probe sets are OsAffx.26632.2.S1_x_at and Os.39997.2.S1_a_at, and are described as Nucleolar genes and retrotransposons respectively (Table S2).

Probe set transcription profiles between plants inoculated with parental line Sc2 or segregated lines Sc2a, Sc2b, Sc2c, Sc2d, Sc2e or Sc2f compared to uninoculated plants.

We were able to detect 235 probe sets that were differentially transcribed between plants inoculated with Sc2 or with one of the associated segregated lines and uninoculated plants. Among these 235 probe sets, 195 were up-regulated and 40 were down-regulated in

inoculated plants. We carried out GO analysis on these two groups of probe sets. We detected no significant pathways that were differentially transcribed in both groups. A global GO analysis of the 235 probe sets, that were differentially transcribed between plants inoculated with Sc2 or with one of the associated segregated lines and plants uninoculated did not allow us to identify differentially transcribed pathways.

GO	Type	Functions	Count	p-value	FDR
GO:0005976	BP	polysaccharide metabolic process	5	8.00E-06	**0.00036**
GO:0005975	BP	carbohydrate metabolic process	6	0.0001	**0.0024**
GO:0006950	BP	response to stress	6	0.00039	**0.0059**
GO:0050896	BP	response to stimulus	7	0.001	**0.012**
GO:0031988	CC	membrane-bounded vesicle	17	0.0039	**0.054**
GO:0031982	CC	vesicle	17	0.004	**0.054**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	17	0.0039	**0.054**
GO:0031410	CC	cytoplasmic vesicle	17	0.004	**0.054**

Table 1: Significant GO terms for genes commonly altered in their transcription in plants inoculated with an AMF lines compared to the uninoculated treatment.

GO: number of GO terms, Type: BP: Biological process; MF: Molecular function; CC: Cell cycle; count: number of altered probe sets found in this GO term; P-value is given for Fisher test. FDR: False discovery Rate. (adjusted p-value). In bold are present the significant values (≥0.1 for the FDR).

Probe set transcription profiles between plants inoculated with parental line S3 or segregated lines S3a, S3b or S3c compared to uninoculated plants.

We identified 333 probe sets that were differentially transcribed between plants inoculated with parental line S3 or with one of the associated segregated lines and uninoculated plants. 269 of these 333 probe sets were down regulated, 59 were up-regulated in inoculated plants and 5 probe sets were up-regulated in plants inoculated with lines S3a and S3c and down-regulated in the plants inoculated with S3 and S3b. These five probe sets correspond to five genes that have no known function (Table S3). The GO ontology analysis carried out on the 59 up-regulated probe sets and on the 269 down-regulated probe sets allowed us to identify pathways specifically affected only for the group of the 269 up-regulated probe sets. The only significant GO term was involved in carbohydrate metabolic processes (Table S4).

Probe set transcription profiles between plants inoculated with parental line S4 or segregated lines S4a, S4b or S4c compared to uninoculated plants.

We found that 214 probe sets were differentially transcribed between plants inoculated with parental line S4 or with one of the associated segregated lines and uninoculated plants. Among these 214 probe sets, 99 probe sets were found to be up-regulated and 115 were found to be down regulated. GO analysis on the 99 up-regulated probe sets revealed that pathways involved in carbohydrate biosynthetic processes and in carbohydrate metabolic processes were up-regulated (Table S5). Results of the GO analysis on the 115 probe sets have shown that pathways involved in response to stress, stimulus and that hydrolase activities were down-regulated (Table S6).

Probe sets	Function
Down regulated	
Os.21854.2.S1_at	Expressed protein
Os.13846.2.S1_at	Indole-3-acetic acid-amido synthetase GH3.7, putative, expressed
Os.9138.1.S1_at	RING-H2 finger protein ATL2A, putative, expressed
OsAffx.3920.1.S1_at	Terpene synthase 7, putative, expressed
Os.50510.1.S1_at	Expressed protein
Os.55743.1.S1_at	Expressed protein
Os.9194.1.S1_at	Phi-1-like phosphate-induced protein, putative, expressed
Os.49208.1.S1_at	Serine/threonine-protein kinase receptor precursor, putative, expressed
Up regulated	
Os.53402.1.S1_at	Ent-kaurene synthase B, chloroplast precursor, putative, expressed
Os.17112.1.S1_at	Leucoanthocyanidin dioxygenase, putative, expressed
Os.16305.1.S1_at	Cytochrome P450 CYP99A1, putative, expressed
Os.30746.1.S1_at	Cytochrome P450 71D10, putative, expressed
OsAffx.28405.1.S1_at	Terpene synthase 7, putative
Os.22681.1.S1_at	Cytochrome P450 72A1, putative, expressed
OsAffx.28405.2.S1_x_at	Terpene synthase 7, putative, expressed
Os.51214.2.S1_at	Expressed protein
Os.51496.1.S1_at	Pathogenesis-related protein PRB1-2 precursor, putative, expressed
Os.38113.2.S1_at	Cytochrome P450 71D10, putative, expressed
Os.33370.1.S1_at	Potassium channel KAT1, putative, expressed

Table 2: Probe sets that were commonly up-regulated and down-regulated in plants inoculated with S2a, Sc2b, Sc2c, Sc2d, Sc2e and Sc2f compared to plants inoculated with the parental line Sc2.

Differences observed in transcription profiles between plants inoculated with Sc2, Sc2a, Sc2b, Sc2c, Sc2d, Sc2e or Sc2f

We identified 19 probe sets that were commonly differentially transcribed between the plants inoculated with the parental line Sc2 and the associated segregated lines. Eight of them were down-regulated and eleven were up-regulated in the plants inoculated with the segregated lines (Table 2). These probe sets belonged to different protein encoding families. Four probe sets among the 11 up-regulated were genes encoding cytochrome P450 enzymes. GO analysis did not permit us to identify any specific pathways concerning these 19 probe sets.

	Sc2a	Sc2b	Sc2c	Sc2d	Sc2e	Sc2f
Sc2	353 / 223	338 / 220	389 / 531	123 / 233	464 / 359	363 / 460
Sc2a	-	302 / 250	360 / 547	267 / 354	464 / 588	381 / 498
Sc2b	-	-	206 / 524	107 / 325	446 / 574	293 / 497
Sc2c	-	-	-	258 / 293	625 / 979	683 / 449
Sc2d	-	-	-	-	601 / 371	413 / 258
Sc2e	-	-	-	-	-	486 / 554
Sc2f	-	-	-	-	-	-

	S3a	S3b	S3c
S3	190 / 179	172 / 195	506 / 219
S3a	-	111 / 122	481 / 150
S3b	-	-	549 / 222
S3c	-	-	-

	S4a	S4b	S4c
S4	205 / 179	213 / 208	167 / 261
S4a	-	266 / 289	292 / 330
S4b	-	-	419 / 314
S4c	-	-	-

	Sc2 vs S4	Sc2 vs S3	S4 vs S3
Genes number	591/100	441/287	175/590

Table 3: Number of probe set differentially transcribed between the plants inoculated with the different treatments. In each case, first number represents the number of down-regulated probe sets, and the second number represents the number of up-regulated probe sets.

GO	Type	Functions	Count	p-value	FDR
GO:0016998	BP	cell wall macromolecule catabolic process	7	1.90E-08	**8.60E-06**
GO:0006032	BP	chitin catabolic process	5	1.30E-07	**3.00E-05**
GO:0006026	BP	aminoglycan catabolic process	5	6.10E-07	**6.90E-05**
GO:0006030	BP	chitin metabolic process	5	6.10E-07	**6.90E-05**
GO:0044036	BP	cell wall macromolecule metabolic process	7	7.80E-07	**7.00E-05**
GO:0009607	BP	response to biotic stimulus	11	2.20E-06	**0.00017**
GO:0006022	BP	aminoglycan metabolic process	5	1.50E-05	**0.00096**
GO:0051704	BP	multi-organism process	11	3.00E-05	**0.0017**
GO:0051707	BP	response to other organism	9	4.30E-05	**0.0022**
GO:0000272	BP	polysaccharide catabolic process	5	0.00032	**0.014**
GO:0050896	BP	response to stimulus	30	0.00041	**0.017**
GO:0006952	BP	defense response	9	0.0011	**0.039**
GO:0006950	BP	response to stress	20	0.0011	**0.039**
GO:0006811	BP	ion transport	12	0.0038	0.12
GO:0006720	BP	isoprenoid metabolic process	5	0.0042	0.13
GO:0005976	BP	polysaccharide metabolic process	8	0.0047	0.13
GO:0030001	BP	metal ion transport	7	0.0055	0.15
GO:0006812	BP	cation transport	9	0.019	0.45
GO:0019748	BP	secondary metabolic process	7	0.019	0.45
GO:0006575	BP	cellular aminoacid derivative metabolic process	7	0.049	0.96
GO:0004568	MF	chitinase activity	7	5.90E-07	**0.00013**
GO:0004497	MF	monooxygenase activity	10	4.00E-05	**0.0029**
GO:0016491	MF	oxidoreductase activity	28	3.50E-05	**0.0029**
GO:0004553	MF	hydrolase activity	10	0.0001	**0.0043**
GO:0016798	MF	hydrolase activity, acting on glycosyl bonds	11	9.20E-05	**0.0043**
GO:0005507	MF	copper ion binding	5	0.0012	**0.044**
GO:0020037	MF	heme binding	7	0.0054	0.15
GO:0005506	MF	iron ion binding	11	0.0055	0.15
GO:0030246	MF	carbohydrate binding	5	0.0062	0.15
GO:0046906	MF	tetrapyrrole binding	7	0.0073	0.16
GO:0005215	MF	transporter activity	18	0.0082	0.16
GO:0004386	MF	helicase activity	6	0.02	0.35
GO:0031982	CC	vesicle	140	7.50E-10	**1.70E-08**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	140	7.00E-10	**1.70E-08**
GO:0031988	CC	membrane-bounded vesicle	140	7.00E-10	**1.70E-08**
GO:0031410	CC	cytoplasmic vesicle	140	7.50E-10	**1.70E-08**

Table 4: significant GO term for genes down-regulated in plants inoculated with Sc2c isolate compared to the parental isolate Sc2. Column heading follow Table 1.

Then, we compared each probe set profile of plants inoculated with segregated lines separately with probe set profiles of plant inoculated with the parental line Sc2. Numbers of differentially probe sets are listed in Table 3. Results of GO analysis on each comparison revealed that pathways involved in polysaccharide or carbohydrate metabolism were up-regulated and down-regulated in their transcription in comparison in plants inoculated with Sc2c or Sc2f and with plants inoculated with the parental line Sc2 respectively (Tables 4 and

5 respectively). Pathways involved in chitin metabolic process, hydrolase activity and responses to different stimuli and stresses were up-regulated in plants inoculated with Sc2c and down-regulated in plants inoculated with Sc2f and Sc2b compared to plants inoculated with parental line Sc2 (Table 4, 5 and S7 respectively).

GO	Type	Functions	Count	p-value	FDR
GO:0006950	BP	response to stress	19	2.60E-05	**0.0088**
GO:0034641	BP	cellular nitrogen compound metabolic process	12	0.00012	**0.014**
GO:0009266	BP	response to temperature stimulus	6	9.00E-05	**0.014**
GO:0000271	BP	polysaccharide biosynthetic process	5	0.00027	**0.022**
GO:0050896	BP	response to stimulus	23	0.0004	**0.027**
GO:0009628	BP	response to abiotic stimulus	9	0.0018	**0.097**
GO:0016051	BP	carbohydrate biosynthetic process	7	0.002	**0.097**
GO:0034637	BP	cellular carbohydrate biosynthetic process	6	0.0025	0.11
GO:0019318	BP	hexose metabolic process	6	0.0051	0.19
GO:0006006	BP	glucose metabolic process	5	0.0065	0.22
GO:0044271	BP	cellular nitrogen compound biosynthetic process	7	0.0074	0.23
GO:0006952	BP	defense response	6	0.008	0.23
GO:0044262	BP	cellular carbohydrate metabolic process	8	0.0089	0.23
GO:0005996	BP	monosaccharide metabolic process	6	0.01	0.24
GO:0005975	BP	carbohydrate metabolic process	10	0.024	0.52
GO:0006066	BP	alcohol metabolic process	6	0.023	0.52
GO:0005976	BP	polysaccharide metabolic process	5	0.032	0.64

Table 5: significant GO term for genes up-regulated in plants inoculated with Sc2f isolate compared to the parental isolate Sc2. Column heading follow Table 1.

Finally, we compared the probe sets profiles of plants inoculated with segregated lines with each other. Table 3 presents the number of probe sets differentially transcribed. GO analysis on each group of probe sets allowed us to detect that different pathways are altered. Carbon metabolism, carbon storage and carbon utilization are some of the pathways that were differentially transcribed among plants inoculated with different segregated lines as we can see for plants inoculated with Sc2b and plants inoculated with Sc2a (Table 6), or plants inoculated with Sc2e and plants inoculated with Sc2c (Table S8). Pathways involved in chitinase processes, response to stress and others stimulus were shown to be altered in almost all comparisons. Additionally, pathways involving various organic and inorganic nitrogenous compounds were altered in their transcription in the comparison among plants inoculated with segregated lines Sc2f and Sc2c, Sc2f and Sc2e, Sc2b and Sc2a, Sc2e and Sc2a and Sc2f and Sc2b (Table S9-S13 respectively).

GO	Type	Functions	Count	p-value	FDR
GO:0006090	BP	pyruvate metabolic process	8	2.60E-06	**0.00075**
GO:0006950	BP	response to stress	17	2.10E-05	**0.0031**
GO:0009790	BP	embryonic development	6	0.00019	**0.019**
GO:0050896	BP	response to stimulus	20	0.00041	**0.03**
GO:0015976	BP	carbon utilization	5	0.00052	**0.03**
GO:0019318	BP	hexose metabolic process	6	0.0018	**0.077**
GO:0032787	BP	monocarboxylic acid metabolic process	9	0.0018	**0.077**
GO:0005996	BP	monosaccharide metabolic process	6	0.0038	0.14
GO:0006091	BP	generation of precursor metabolites and energy	7	0.0046	0.15
GO:0007275	BP	multicellular organismal development	9	0.0084	0.21
GO:0044262	BP	cellular carbohydrate metabolic process	7	0.0096	0.21
GO:0032502	BP	developmental process	10	0.008	0.21
GO:0006066	BP	alcohol metabolic process	6	0.0093	0.21
GO:0032501	BP	multicellular organismal process	9	0.012	0.24
GO:0016051	BP	carbohydrate biosynthetic process	5	0.014	0.28
GO:0009628	BP	response to abiotic stimulus	6	0.023	0.42
GO:0044248	BP	cellular catabolic process	8	0.028	0.48
GO:0006732	BP	coenzyme metabolic process	5	0.032	0.52
GO:0005975	BP	carbohydrate metabolic process	8	0.044	0.67
GO:0044265	BP	cellular macromolecule catabolic process	5	0.047	0.68
GO:0005199	MF	structural constituent of cell wall	5	0.00058	**0.029**
GO:0016616	MF	oxidoreductase activity	5	0.00081	**0.029**
GO:0016614	MF	oxidoreductase activity	5	0.0013	**0.031**
GO:0004175	MF	endopeptidase activity	7	0.018	0.33
GO:0008233	MF	peptidase activity	9	0.027	0.39
GO:0005198	MF	structural molecule activity	8	0.043	0.52

Table 6: significant GO term for genes down-regulated in plants inoculated with Sc2b isolate when compare to plants inoculated with Sc2a isolate. Column heading follow Table 1.

Differences observed in transcription profiles between plants inoculated with S3, S3a, S3b or S3c

Seventy-two probe sets were found to be differentially transcribed in common between plants inoculated with parental line S3 and the plants inoculated with the associated segregated lines. Among these 72 probe sets, 56 were down-regulated and 16 were up-regulated in plants inoculated with segregated lines. The GO analysis on the 56 down-regulated probe sets and on the 16 up-regulated probe sets did not give significant results. Nevertheless, the GO analysis on the total differentially probe sets allowed us to detect that pathways involved in response to stimulus, transport, and ion binding were altered (Table 7).

GO	Type	Functions	Count	p-value	FDR
GO:0050896	BP	response to stimulus	6	0.0056	**0.064**
GO:0051234	BP	establishment of localization	5	0.031	**0.093**
GO:0006810	BP	transport	5	0.031	**0.093**
GO:0051179	BP	localization	5	0.033	**0.093**
GO:0020037	MF	heme binding	5	3.60E-06	**6.90E-05**
GO:0046906	MF	tetrapyrrole binding	5	4.70E-06	**6.90E-05**
GO:0005506	MF	iron ion binding	5	0.00012	**0.0012**
GO:0043169	MF	cation binding	9	0.0026	**0.013**
GO:0046872	MF	metal ion binding	9	0.002	**0.013**
GO:0043167	MF	ion binding	9	0.0026	**0.013**
GO:0046914	MF	transition metal ion binding	6	0.015	**0.062**
GO:0031988	CC	membrane-bounded vesicle	18	0.0027	**0.037**
GO:0031982	CC	vesicle	18	0.0028	**0.037**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	18	0.0027	**0.037**
GO:0031410	CC	cytoplasmic vesicle	18	0.0028	**0.037**

Table 7: significant GO term for genes altered in their transcription in plant inoculated with S3a, S3b and S3c isolate compared to plants inoculated with parental S3 isolate. Column heading follow Table 1.

One probe set (Os.26226.1.S1) was up-regulated in plants inoculated with segregated line S3c and down-regulated in plants inoculated with segregated lines S3a and S3b in comparison with probe set profiles for the plants inoculated with the parental line S3 (Table S14). This probe set belonged to a gene coding for an ATP binding protein. Two other probe sets were detected as commonly altered between plants inoculated with the parental line S3 and the plants inoculated with the associated segregated lines. These probe sets were up-regulated in plants inoculated with S3a and S3b but down-regulated in plants inoculated with S3c (Table S14). Os.8655.1.S1 corresponds to a gene coding for thionin precursor, and Os.54146.1.S1 corresponds to a gene coding for a nitrate reductase.

Comparisons between plants inoculated with the parental line S3 and each of the segregated lines S3a, S3b and S3c separately were performed. Numbers of probe sets altered are shown in Table 3. As found with the previous segregated lines, GO analysis results revealed that pathways involved in stress and response to stimuli were altered in their transcription for the three comparisons (Table S15-S17). Moreover, ion-binding processes, lipid metabolism and carbohydrate metabolism were altered too.

GO	Type	Functions	Count	p-value	FDR
GO:0051179	BP	localization	15	0.00084	**0.023**
GO:0050896	BP	response to stimulus	14	0.0008	**0.023**
GO:0006950	BP	response to stress	11	0.00034	**0.023**
GO:0006810	BP	transport	14	0.0019	**0.032**
GO:0051234	BP	establishment of localization	14	0.002	**0.032**
GO:0016491	MF	oxidoreductase activity	11	0.0025	**0.079**
GO:0016787	MF	hydrolase activity	19	0.0022	**0.079**
GO:0005506	MF	iron ion binding	6	0.0041	**0.087**
GO:0003824	MF	catalytic activity	43	0.011	0.17
GO:0008233	MF	peptidase activity	7	0.016	0.2
GO:0070011	MF	peptidase activity	6	0.025	0.26
GO:0043167	MF	ion binding	15	0.038	0.3
GO:0043169	MF	cation binding	15	0.038	0.3
GO:0022857	MF	transmembrane transporter activity	5	0.05	0.35
GO:0031982	CC	vesicle	36	0.036	0.53
GO:0016023	CC	cytoplasmic membrane	36	0.036	0.53
GO:0031410	CC	cytoplasmic vesicle	36	0.036	0.53
GO:0031988	CC	membrane-bounded vesicle	36	0.036	0.53

Table 8: significant GO term for genes up-regulated in plants inoculated with S4a isolate compared to plants inoculated with parental isolate S4. Column heading follow Table 1.

Comparison between plants inoculated with segregated lines and GO analysis were carried out. Numbers of probe sets that were altered are shown in Table 3. No pathway was significantly altered in the comparison between plants inoculated with S3b and S3a. Alteration in pathways involved in ion binding was observed for the comparison between plants inoculated with S3c and S3a, and between plants inoculated with S3c and S3b (Table S18 and S19).

Differences observed in transcription profiles between plants inoculated with S4, S4a, S4b or S4c

Ten probe sets were found to be commonly down-regulated and one probe set was found to be commonly up-regulated in the comparison between the probe set profiles of plants inoculated with the parental line S4 and the probe set profiles of the plants inoculated with the associated segregated lines. These probe sets are listed in Table S20. All the genes corresponding to these probes sets code for unknown proteins except for two, that code for a reductase and xylanase.

We separately compared each probe set profile of plants inoculated with segregated lines S4a, S4b and S4c with the probe set profiles of plants inoculated with parental line S4. Numbers of probe sets altered are shown in Table 3. GO analysis results showed that terpenoid pathways were altered in comparison to plants inoculated with S4b with the plants inoculated with S4 (Table S21). Transport pathways were up-regulated in comparison to plants inoculated with S4a with the plants inoculated with S4 (Table 8). These transport pathways concerned macromolecules, small molecules and ions. One of the 14 probe sets that we found in our treatment (Os.21638.1.S1) belonging to this GO term was Phosphate transporter 11 (Pt11), an inorganic phosphate transporter gene that is known to respond specifically to colonization by AMF (Paszkowski *et al.* 2002).

We compared the transcription profiles of the plants inoculated with segregated lines with each other. Numbers of probe sets that were differentially transcribed are shown in Table 3. GO analysis was carried out on each group of probe sets. Results showed that pathways involved in terpenoid metabolism were altered between the plants inoculated with S4c and S4b as well as pathways involved in chitin metabolism and catabolic processes (Table S22). Pathways playing a role in polysaccharide metabolism processes were differentially transcribed between plants inoculated with segregated lines S4c and S4a (Table S23)

Differences in probe set profiles between plants inoculated with parental lines and their associated segregated lines and probe set profiles between parental lines.

We compared each probe set profile of plants inoculated with parental lines with each other. Numbers of probe sets altered are shown in Table 3. GO analysis were carried out on each group of probe sets. GO analysis did not permit us to identify enrichment in specific metabolic pathways in comparison between plants inoculated with Sc2 and plants inoculated with S4. GO analysis for the probe sets commonly altered between plants inoculated with parental lines S3 and Sc2, and plants inoculated with parental lines S4 and S3 are presented in Table S24 and S25 respectively. Results showed that pathways involved in response to stress, ion bindin, hydrolase, chitinase activities and carbohydrate metabolism were found to be altered in their transcription.

Discussion

In this study, we detected which genes were altered in their transcription by the presence of an AMF isolate. Comparison between transcription profiles of plants inoculated with segregated lines allowed us to demonstrate that rice transcriptome can be altered by segregation that occurred in AMF.

In 2005, Guimil *et al.* (2005) carried out a comparison of data generated from microarrays hybridized with RNA coming from rice roots inoculated with AMF or uninoculated. 224 genes were found to be specifically transcribed in the presence of AMF. We investigated whether these 224 genes were altered in our experiment. Among the 47 probe sets found to be differentially transcribed between plants inoculated with AMF and plants uninoculated, only three probe sets matched with the data from Guimil *et al.* (2005) genes list (Table S26). These three genes were found to be up-regulated in our experiment whereas two were up-regulated (AM84 and AM189) and one down-regulated (AM213) in the study by Guimil *et al.* (2005). AM213 was also found, in chapter 2, to be up-regulated in the comparison between plants inoculated with AMF and uninoculated plants. One explanation of these differences between these experiments could be tha different AMF isolates were used in these two studies. In Chapter 2, we showed that plant gene transcription can be different between AMF isolates. Here, we carried out the analysis after pooling data on plants inoculated with fifteen genetically different isolates and which were genetically distant from the isolate used by Guimil *et al.* (2005) (DAOM 181602). Moreover, the RNA samples were extracted at 6 weeks after inoculation in the study by Guimil *et al.* (2005), whereas we performed the RNA extractions 12 weeks after inoculation. This shift in the timing of the RNA extraction may explain the differences in transcriptome profiles obtained between our experiment and that of Guimil *et al.* (2005).

Effects of segregation on plant gene transcription

In 2010, Angelard *et al.* (2010) investigated the effects of segregation in AMF on the transcription of four AM specific plant genes using a quantitative PCR approach. In their experiment, they used the parental line S4 and the segregated lines S4a, S4b and S4c. They found that segregation in AMF could alter the transcription of AM specific plant genes. We found the same trend for AM3, AM14 and Pt11. AM1 was not detected as differentially

transcribed between the plants inoculated with these isolates. One explanation could be the fact that the microarray technique is less sensitive than quantitative PCR. The advantage of the microarray approach was that we were able to detect numerous plant genes that were altered in their transcription. Nevertheless, further investigations have to be performed on specific genes.

Defence

Interactions between organisms can either be beneficial, such as mutualistic symbioses, or negative, such as parasitism. In 2005, Guimil *et al.* compared the rice transcriptome when the plant was inoculated with AMF (*Glomus intraradices*) or uninoculated, but also with rice inoculated with pathogens *Magnaporthe grisea* and *Fusarium moniliforme*. They found that some genes are commonly differentially transcribed no matter whether the partner was a mycorrhizal fungus or a pathogenic fungus. To form the AM symbiosis, fungi need to be recognized by the host plant. Evidence suggests that, as in plant-pathogen interactions, the suppression or activation of plant defence mechanisms play a key role in AM fungal colonization (Garcia-Garrido *et al.* 2002; Kloppholz *et al.* 2011)) even if a recent study show that the activation of basal defence mechanisms of rice by AMF did not affect the AM symbiosis (Campos-Soriano *et al.* 2010). It is thought that hydrolytic enzymes (such as hydrolases or chitinase) could hydrolyse fungal products (called elicitors) that are recognized by the plant as signals of potential aggressor (Garcia-Garrido *et al.* 2002). Moreover, Salzer and Boller (Salzer *et al.* 2000) proposed that constitutively expressed plant chitinases are enzymes responsible for the degradation of the elicitor. Our results show that many of the GO terms that were altered by segregation in AMF contained genes that are involved in defence or stress. Our results showed that pathways involved in plant defence, stress, chitinase or other hydrolase activities were altered when we compared the effect on plants of the different segregated lines between each other, or between segregated lines and respective parental lines. For example, we found that GO terms for chitinases, or responses to stimuli were altered when we compared the effects of segregated lines S4c versus S4b (Table S22), S3c versus S3b (Table S19) or Sc2b and Sc2a (Table14). These results show that segregation in AMF could induce a different transcriptomic response in the plant.

Inorganic Phosphate

A main role of the fungus is to supply phosphate to the plant. We found almost no differences in pathways involved in phosphate transport in our study, except in the comparison between plants inoculated with the parental line S4 and the respective segregated lines S4a, S4b and S4c. Our results showed that the AM specific gene Pt11 was differentially transcribed between plants inoculated with segregated lines S4a and S4c compared to the plants inoculated with the parental line S4 and the segregated line S4b. Guimil *et al.* (2005) found that Pt11 transcription was increased, but they extracted the RNA at 6 weeks after inoculation. In our experiment, we extracted the RNA at 12 weeks after inoculation, when no difference in plant growth was detectable. In chapter 2, we did not detect transcriptional differences in this gene after 12 weeks of plant growth in their experiment. These results suggest that segregation in AMF could potentially alter nutrient exchange between plants and AMF and, therefore, showed that given isolate might be specifically altered plant gene transcription.

Intra- versus inter-specific effects of AMF on plants

We compared altered pathways found in comparison between parental isolate with each other and pathways found in comparison between segregated lines and parental lines. Pathways involved in response to stress, chitinase activities and ion binding are in common in many of these comparisons. This result suggested that segregated lines are not more related to their parental line than the parental lines are related from each other. It revealed the power of the segregation in AMF to create variation in symbiotic effects in a short time.

Conclusion

Overall, our results show that segregation in AMF has an impact on rice transcriptome profiles. We showed that segregation can act on genes that could be involved in nutrient exchange between the two hosts, but also plays a role in the recognition of the fungi by the plant. Our study contributes to understanding how much the genetics of mycorrhizal fungi influences plant gene transcription.

Acknowledgments

This work was supported by grants 31000AO-105790/1 and 31003A-127371 from the Swiss National Science Foundation. We thank U. Paszkowski for providing rice seeds, J. Bonvin for assistance with fungal inoculation and the rest of our group for their helpful comments on the manuscript. We also thank the Genomic Technology Facility of the University of Lausanne for their assistance for the microarray and quantitative PCR.

Supplementary files

Microarray validation

We used several different approaches to check the validity of the transcription profiles generated from the microarrays. These were:

1) Quantitative PCR

Quantitative PCR was carried out on selected genes. New RNA extractions were performed on the original samples. We used the RNeasy Plant mini kit (Qiagen Inc®) to extract RNA following the manufacturer's recommendations. Precipitation with LiCl was also performed (Ambion®)

Quantitative PCRs were performed on three genes that were detected in the microarrays because of their fold change expression between the isolates and the NM treatment: Os.52079.1.S1 (unknown function, but contain a domain mainly found in plants), OS.7947.1.S1 (Lichenase-2 precursor) and Os.5319.1.S1 (Phosphoglucomutase, and play a role in carbohydrate metabolism). Primer 3 software was used to design the primers. PCR conditions and mixs were adopted from Gutjahr *et al*. (2008).

Power Sybr® green was used as the label (Applied Biosystems) for quantitative PCR and each sample was run in triplicate. Quantitative PCR was performed on Prism AB7900 quantitative PCR machine. Ubiquitin transcript levels were used to normalize the data. Transcription values were calculated with SDS 2.3 software (Applied Biosystems) and qBASE Plus software (Hellemans *et al.* 2007).

Quantitative PCR confirmed the microarrays results (Figure S1, Table S1). For example, Os.52079.1.S1 was down-regulated by a fold change of 11 and 10 for C2 and D1 respectively on microarray. Quantitative PCR showed a similar fold change and differences were statistically different. Quantitative PCR results on Os.5319.1.S1 transcript level were congruent with the microarray results although the differences were not significant.

Analyses on gene transcription levels were performed for each gene separately. When ANOVAs were significant, we carried-out a multiple comparison test using Tukey-kramer HSD test with an alpha level of 0.05.

All analyses were performed with the statistical program JMP® version 5.0 (SAS Institute Inc., Cary, NC, USA).

2) Gene Ontology analysis

We made gene Ontology (GO) analysis on the microarray results with the software AgriGO (website: http://bioinfo.cau.edu.cn/agrigo/). The GO analyses were performed on the probe sets with fold changes 2 or more or -2 or less. The GO terms generated in the GO analysis tests whether a significant number of genes that are up or down-regulated group together into a given metabolic pathways, cellular functions or the cell cycle. If the microarray data were not reliable then significant GO terms should not be generated in the analysis. The significance of the GO terms is determined by a false discovery rate (FDR) statistic. Many of the significant GO terms in our analysis also contained a large number of significantly up or down-regulated probe sets, which would not be expected if the microarray data were simply showing random patterns. The GO analysis also reveals positive and negative regulation among significant GO terms. Again significant positive and negative relationship between GO terms would not be expected.

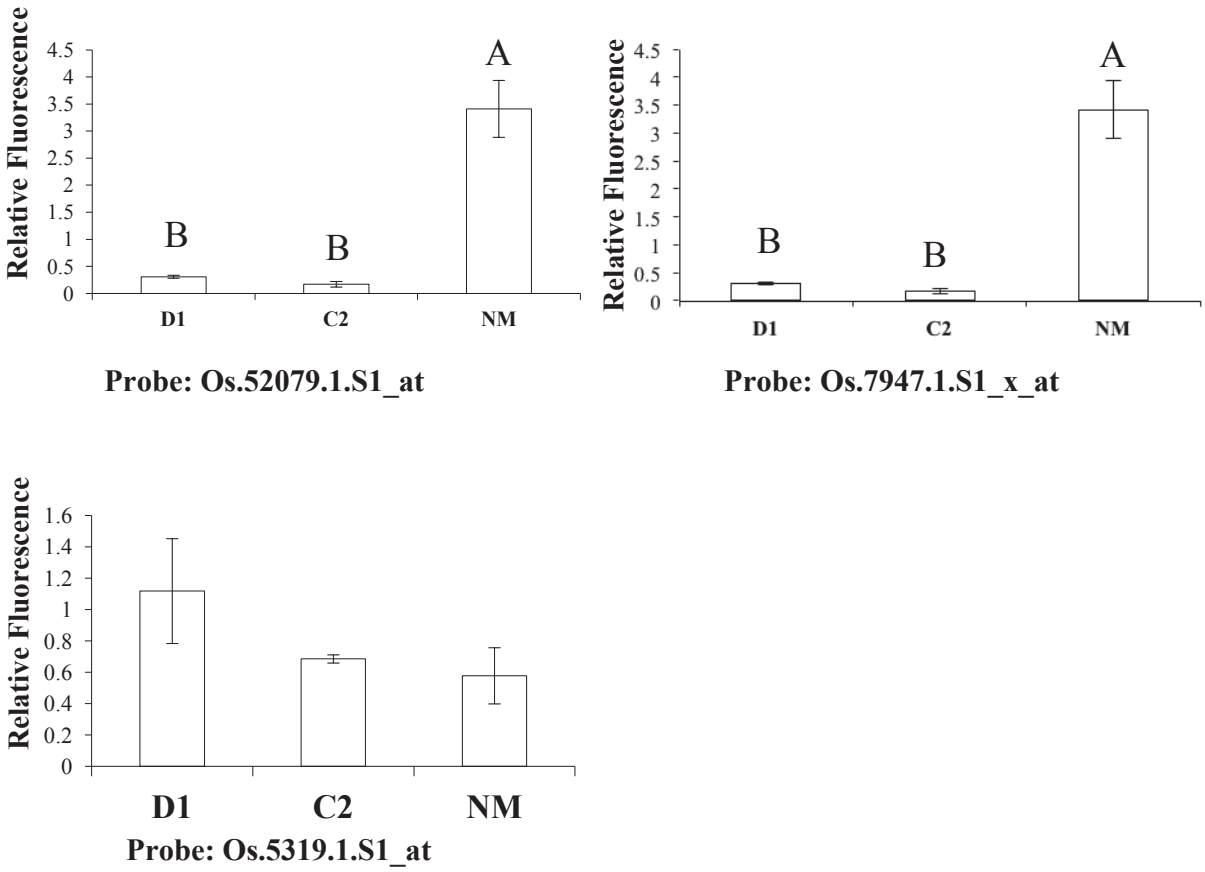


Figure S1: Quantitative PCR for the validation of microarray using three probe sets.

Error bars ± 1 SE. Bars topped by different letters are significantly different at $p \leq 0.05$.

Source	d.f.	SS	F	*p*	Same trends as in microarray data
Os.52079.1.S1_at	**5**	**16.034**	**7.403**	**0.0022***	**Yes**
Os.7947.1.S1_x_at	**5**	**12.293**	**59.813**	**<0.001***	**Yes**
Os.5319.1.S1_at	**5**	**2.126**	**2.323**	**0.1075**	**Yes**

Table S1: Analysis of variance on plant transcription of three genes for the microarray validation.

Probe sets	Putative functions
OsAffx.26632.2.S1_x_at	Nucleolar, Nop52 family protein
Os.39997.2.S1_a_at	Retrotransposon protein

Table S2: Probe sets that were up-regulated in plants inoculated with S4, S4a, S4b and S4c isolate and down-regulated in plants inoculated with the other AMF isolates compared to the uninoculated treatment

Probe sets	Putative function
OsAffx.22999.1.S1_at	Conserved hypothetical protein
Os.10659.1.S1_s_at	Conserved hypothetical protein
OsAffx.20082.1.S1_at	Conserved hypothetical protein
Os.43929.1.S1_s_at	Conserved hypothetical protein
Os.8823.1.S1_at	Conserved hypothetical protein

Table S3: Probe sets that were up-regulated in plants inoculated with S3a and S3b isolates and down-regulated in plants inoculated with S3 and S3c isolates compared to the uninoculated treatment.

GO	Type	Functions	Count	p-value	FDR
GO:0005975	BP	carbohydrate metabolic process	6	0.00029	**0.012**
GO:0006950	BP	response to stress	5	0.0056	0.11
GO:0050896	BP	response to stimulus	6	0.011	0.15
GO:0043167	MF	ion binding	7	0.046	0.23
GO:0043169	MF	cation binding	7	0.046	0.23
GO:0046872	MF	metal ion binding	7	0.039	0.23
GO:0003824	MF	catalytic activity	14	0.17	0.49
GO:0016740	MF	transferase activity	7	0.15	0.49
GO:0005488	MF	binding	12	0.38	0.92
GO:0009507	CC	chloroplast	5	0.0046	0.29

Table S4: Significant GO terms for genes up-regulated in plants inoculated with isolates S3, S3a, S3b and S3c compared to the uninoculated treatment. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0034637	BP	cellular carbohydrate biosynthetic process	5	5.80E-05	**0.0053**
GO:0016051	BP	carbohydrate biosynthetic process	5	0.00019	**0.0086**
GO:0005975	BP	carbohydrate metabolic process	7	0.00068	**0.021**
GO:0044262	BP	cellular carbohydrate metabolic process	5	0.0018	**0.04**
GO:0032787	BP	monocarboxylic acid metabolic process	5	0.0033	**0.059**
GO:0019752	BP	carboxylic acid metabolic process	8	0.0084	**0.094**
GO:0043436	BP	oxoacid metabolic process	8	0.0084	**0.094**
GO:0042180	BP	cellular ketone metabolic process	8	0.0093	**0.094**
GO:0006082	BP	organic acid metabolic process	8	0.0084	**0.094**
GO:0016491	MF	oxidoreductase activity	10	0.00019	**0.0048**
GO:0003824	MF	catalytic activity	26	0.033	**0.42**
GO:0031982	CC	vesicle	30	0.0012	**0.017**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	30	0.0012	**0.017**
GO:0031410	CC	cytoplasmic vesicle	30	0.0012	**0.017**
GO:0031988	CC	membrane-bounded vesicle	30	0.0012	**0.017**

Table S5: Significant GO terms for genes up-regulated in plants inoculated with isolates S4, S4a, S4b and S4c compared to the uninoculated treatment. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0006950	BP	response to stress	9	0.00017	**0.012**
GO:0050896	BP	response to stimulus	9	0.0052	0.18
GO:0004553	MF	hydrolase activity	5	0.00012	**0.0018**
GO:0016798	MF	hydrolase activity	5	0.00028	**0.002**
GO:0005215	MF	transporter activity	6	0.016	**0.077**
GO:0031982	CC	vesicle	38	1.10E-05	**0.00018**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	38	1.00E-05	**0.00018**
GO:0031410	CC	cytoplasmic vesicle	38	1.10E-05	**0.00018**
GO:0031988	CC	membrane-bounded vesicle	38	1.00E-05	**0.00018**
GO:0016020	CC	membrane	16	0.069	0.93

Table S6: Significant GO terms for genes down-regulated in plants inoculated with isolates S4, S4a, S4b and S4c compared to the uninoculated treatment. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0015671	BP	oxygen transport	5	6.40E-11	**1.70E-08**
GO:0015669	BP	gas transport	5	2.20E-10	**3.00E-08**
GO:0006950	BP	response to stress	20	2.50E-06	**0.00023**
GO:0050896	BP	response to stimulus	23	0.00013	**0.0091**
GO:0019825	MF	oxygen binding	5	3.00E-08	**3.20E-06**
GO:0020037	MF	heme binding	6	0.0022	**0.1**
GO:0046906	MF	tetrapyrrole binding	6	0.0029	**0.1**
GO:0005506	MF	iron ion binding	8	0.0073	0.19
GO:0016491	MF	oxidoreductase activity	13	0.035	0.68
GO:0003700	MF	transcription factor activity	8	0.039	0.68
GO:0031982	CC	vesicle	68	0.0051	0.13
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	68	0.005	0.13
GO:0031988	CC	membrane-bounded vesicle	68	0.005	0.13
GO:0031410	CC	cytoplasmic vesicle	68	0.0051	0.13

Table S7: Significant GO terms for genes down-regulated in plants inoculated with isolate Sc2b compared to the parental isolate Sc2. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0032787	BP	monocarboxylic acid metabolic process	26	8.00E-10	**6.20E-07**
GO:0034641	BP	cellular nitrogen compound metabolic process	21	5.60E-07	**0.00022**
GO:0006090	BP	pyruvate metabolic process	11	3.10E-06	**0.0008**
GO:0044262	BP	cellular carbohydrate metabolic process	18	6.40E-06	**0.0012**
GO:0005975	BP	carbohydrate metabolic process	24	1.10E-05	**0.0016**
GO:0043436	BP	oxoacid metabolic process	37	1.80E-05	**0.0017**
GO:0016051	BP	carbohydrate biosynthetic process	13	1.90E-05	**0.0017**
GO:0006082	BP	organic acid metabolic process	37	1.80E-05	**0.0017**
GO:0019752	BP	carboxylic acid metabolic process	37	1.80E-05	**0.0017**
GO:0019318	BP	hexose metabolic process	12	3.00E-05	**0.0021**
GO:0042180	BP	cellular ketone metabolic process	37	2.70E-05	**0.0021**
GO:0006006	BP	glucose metabolic process	10	5.20E-05	**0.0031**
GO:0006066	BP	alcohol metabolic process	14	5.10E-05	**0.0031**
GO:0019320	BP	hexose catabolic process	9	8.80E-05	**0.0045**
GO:0006007	BP	glucose catabolic process	9	8.80E-05	**0.0045**
GO:0005996	BP	monosaccharide metabolic process	12	0.00012	**0.0057**
GO:0044275	BP	cellular carbohydrate catabolic process	9	0.00018	**0.0072**
GO:0046164	BP	alcohol catabolic process	9	0.00017	**0.0072**
GO:0046365	BP	monosaccharide catabolic process	9	0.00017	**0.0072**
GO:0016052	BP	carbohydrate catabolic process	10	0.00025	**0.0096**
GO:0044255	BP	cellular lipid metabolic process	16	0.00044	**0.016**
GO:0050896	BP	response to stimulus	34	0.00058	**0.019**
GO:0006631	BP	fatty acid metabolic process	9	0.00056	**0.019**
GO:0006094	BP	gluconeogenesis	7	0.00055	**0.019**
GO:0006096	BP	glycolysis	7	0.00079	**0.024**
GO:0019319	BP	hexose biosynthetic process	7	0.00087	**0.026**
GO:0046364	BP	monosaccharide biosynthetic process	7	0.001	**0.03**
GO:0046165	BP	alcohol biosynthetic process	7	0.0011	**0.032**
GO:0006633	BP	fatty acid biosynthetic process	7	0.0014	**0.035**
GO:0005976	BP	polysaccharide metabolic process	10	0.0014	**0.035**
GO:0006629	BP	lipid metabolic process	16	0.0019	**0.047**
GO:0009790	BP	embryonic development	7	0.0023	**0.056**
GO:0034637	BP	cellular carbohydrate biosynthetic process	8	0.0031	**0.071**
GO:0008610	BP	lipid biosynthetic process	10	0.0031	**0.071**
GO:0044265	BP	cellular macromolecule catabolic process	11	0.0044	**0.095**
GO:0006950	BP	response to stress	21	0.0044	**0.095**
GO:0009820	BP	alkaloid metabolic process	5	0.0046	**0.097**
GO:0016053	BP	organic acid biosynthetic process	12	0.0058	0.11
GO:0046394	BP	carboxylic acid biosynthetic process	12	0.0058	0.11
GO:0006732	BP	coenzyme metabolic process	10	0.0059	0.11
GO:0006091	BP	generation of precursor metabolites and energy	11	0.0057	0.11
GO:0051186	BP	cofactor metabolic process	11	0.0075	0.14
GO:0042221	BP	response to chemical stimulus	16	0.0082	0.15
GO:0044248	BP	cellular catabolic process	15	0.014	0.24
GO:0015976	BP	carbon utilization	5	0.014	0.24
GO:0033036	BP	macromolecule localization	9	0.014	0.24

GO:0042592	BP	homeostatic process	6	0.016	0.27
GO:0019748	BP	secondary metabolic process	8	0.017	0.28
GO:0044271	BP	cellular nitrogen compound biosynthetic process	9	0.019	0.3
GO:0007242	BP	intracellular signaling cascade	8	0.02	0.31
GO:0019725	BP	cellular homeostasis	5	0.024	0.36
GO:0010033	BP	response to organic substance	9	0.024	0.36
GO:0006073	BP	cellular glucan metabolic process	6	0.03	0.41
GO:0006118	BP	electron transport	9	0.03	0.41
GO:0044042	BP	glucan metabolic process	6	0.03	0.41
GO:0009117	BP	nucleotide metabolic process	6	0.03	0.41
GO:0006753	BP	nucleoside phosphate metabolic process	6	0.03	0.41
GO:0005985	BP	sucrose metabolic process	5	0.04	0.54
GO:0044264	BP	cellular polysaccharide metabolic process	6	0.042	0.54
GO:0007275	BP	multicellular organismal development	13	0.042	0.54
GO:0005984	BP	disaccharide metabolic process	5	0.043	0.55
GO:0016616	MF	oxidoreductase activity	10	4.90E-06	**0.0014**
GO:0016614	MF	oxidoreductase activity	10	1.20E-05	**0.0014**
GO:0016762	MF	xyloglucan:xyloglucosyl transferase activity	5	1.50E-05	**0.0014**
GO:0016491	MF	oxidoreductase activity	32	3.20E-05	**0.0023**
GO:0050662	MF	coenzyme binding	12	0.00022	**0.012**
GO:0016757	MF	transferase activity, transferring glycosyl groups	14	0.001	**0.043**
GO:0020037	MF	heme binding	9	0.0012	**0.043**
GO:0005506	MF	iron ion binding	14	0.0012	**0.043**
GO:0046906	MF	tetrapyrrole binding	9	0.0017	**0.054**
GO:0016758	MF	transferase activity, transferring hexosyl groups	11	0.0022	**0.061**
GO:0048037	MF	cofactor binding	12	0.0025	**0.063**
GO:0050660	MF	FAD binding	5	0.0064	0.15
GO:0043169	MF	cation binding	46	0.012	0.25
GO:0043167	MF	ion binding	46	0.012	0.25
GO:0000166	MF	nucleotide binding	31	0.023	0.4
GO:0046872	MF	metal ion binding	43	0.022	0.4
GO:0031982	CC	vesicle	137	2.90E-05	**0.0011**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	137	2.80E-05	**0.0011**
GO:0031988	CC	membrane-bounded vesicle	137	2.80E-05	**0.0011**
GO:0031410	CC	cytoplasmic vesicle	137	2.90E-05	**0.0011**
GO:0005576	CC	extracellular region	6	0.014	0.39

Table S8: Significant GO terms for genes up-regulated in plants inoculated with isolate Sc2e compared to the parental isolate Sc2c. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0034641	BP	cellular nitrogen compound metabolic process	18	3.00E-07	**0.00014**
GO:0044262	BP	cellular carbohydrate metabolic process	16	1.50E-06	**0.00035**
GO:0006090	BP	pyruvate metabolic process	9	8.80E-06	**0.001**
GO:0005975	BP	carbohydrate metabolic process	20	6.80E-06	**0.001**
GO:0016051	BP	carbohydrate biosynthetic process	10	9.70E-05	**0.0093**
GO:0032787	BP	monocarboxylic acid metabolic process	14	0.00013	**0.0098**
GO:0006066	BP	alcohol metabolic process	11	0.00014	**0.0098**
GO:0019318	BP	hexose metabolic process	9	0.00021	**0.012**
GO:0009790	BP	embryonic development	7	0.00034	**0.018**
GO:0034637	BP	cellular carbohydrate biosynthetic process	8	0.00038	**0.018**
GO:0005996	BP	monosaccharide metabolic process	9	0.00059	**0.023**
GO:0006094	BP	gluconeogenesis	6	0.00054	**0.023**
GO:0019319	BP	hexose biosynthetic process	6	0.00081	**0.028**
GO:0006006	BP	glucose metabolic process	7	0.00079	**0.028**
GO:0050896	BP	response to stimulus	26	0.00098	**0.029**
GO:0046165	BP	alcohol biosynthetic process	6	0.001	**0.029**
GO:0046364	BP	monosaccharide biosynthetic process	6	0.00096	**0.029**
GO:0006950	BP	response to stress	18	0.0013	**0.034**
GO:0043436	BP	oxoacid metabolic process	24	0.0017	**0.038**
GO:0006082	BP	organic acid metabolic process	24	0.0017	**0.038**
GO:0019752	BP	carboxylic acid metabolic process	24	0.0017	**0.038**
GO:0019320	BP	hexose catabolic process	6	0.0019	**0.039**
GO:0006007	BP	glucose catabolic process	6	0.0019	**0.039**
GO:0051186	BP	cofactor metabolic process	10	0.0021	**0.041**
GO:0042180	BP	cellular ketone metabolic process	24	0.0022	**0.041**
GO:0042221	BP	response to chemical stimulus	14	0.0024	**0.044**
GO:0009311	BP	oligosaccharide metabolic process	6	0.0028	**0.048**
GO:0044275	BP	cellular carbohydrate catabolic process	6	0.003	**0.048**
GO:0046164	BP	alcohol catabolic process	6	0.0029	**0.048**
GO:0046365	BP	monosaccharide catabolic process	6	0.0029	**0.048**
GO:0006096	BP	glycolysis	5	0.0044	**0.068**
GO:0006732	BP	coenzyme metabolic process	8	0.007	**0.1**
GO:0016052	BP	carbohydrate catabolic process	6	0.0097	0.14
GO:0019748	BP	secondary metabolic process	7	0.0094	0.14
GO:0010033	BP	response to organic substance	8	0.01	0.14
GO:0044255	BP	cellular lipid metabolic process	10	0.012	0.15
GO:0005984	BP	disaccharide metabolic process	5	0.012	0.16
GO:0016137	BP	glycoside metabolic process	5	0.013	0.16
GO:0006073	BP	cellular glucan metabolic process	5	0.027	0.3
GO:0044042	BP	glucan metabolic process	5	0.027	0.3
GO:0044248	BP	cellular catabolic process	11	0.027	0.3
GO:0005976	BP	polysaccharide metabolic process	6	0.026	0.3
GO:0006629	BP	lipid metabolic process	10	0.028	0.31
GO:0044264	BP	cellular polysaccharide metabolic process	5	0.035	0.38
GO:0006796	BP	phosphate metabolic process	17	0.045	0.45
GO:0006793	BP	phosphorus metabolic process	17	0.045	0.45

GO:0016616	MF	oxidoreductase activity	9	2.40E-06	**0.0005**
GO:0016614	MF	oxidoreductase activity	9	5.60E-06	**0.00059**
GO:0016491	MF	oxidoreductase activity	27	8.70E-06	**0.0006**
GO:0005506	MF	iron ion binding	13	0.00017	**0.0089**
GO:0050662	MF	coenzyme binding	10	0.00022	**0.0093**
GO:0020037	MF	heme binding	8	0.00055	**0.019**
GO:0046906	MF	tetrapyrrole binding	8	0.00079	**0.024**
GO:0048037	MF	cofactor binding	10	0.0019	**0.049**
GO:0043169	MF	cation binding	38	0.0027	**0.057**
GO:0043167	MF	ion binding	38	0.0027	**0.057**
GO:0016791	MF	phosphatase activity	6	0.0034	**0.064**
GO:0042578	MF	phosphoric ester hydrolase activity	6	0.0092	0.16
GO:0046872	MF	metal ion binding	34	0.011	0.18
GO:0016758	MF	transferase activity	7	0.023	0.34
GO:0016757	MF	transferase activity	8	0.037	0.51
GO:0003824	MF	catalytic activity	90	0.045	0.59
GO:0031982	CC	vesicle	96	0.00068	**0.018**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	96	0.00066	**0.018**
GO:0031988	CC	membrane-bounded vesicle	96	0.00066	**0.018**
GO:0031410	CC	cytoplasmic vesicle	96	0.00068	**0.018**

Table S9: Significant GO terms for genes up-regulated in plants inoculated with isolate Sc2f compared to plants inoculated with Sc2c isolate. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0034641	BP	cellular nitrogen compound metabolic process	16	8.90E-06	**0.0042**
GO:0000271	BP	polysaccharide biosynthetic process	5	0.00093	0.15
GO:0005976	BP	polysaccharide metabolic process	9	0.00063	0.15
GO:0005975	BP	carbohydrate metabolic process	15	0.0023	0.27
GO:0009628	BP	response to abiotic stimulus	10	0.0035	0.27
GO:0005996	BP	monosaccharide metabolic process	8	0.003	0.27
GO:0044262	BP	cellular carbohydrate metabolic process	10	0.0055	0.35
GO:0050896	BP	response to stimulus	24	0.006	0.35
GO:0006073	BP	cellular glucan metabolic process	6	0.0082	0.39
GO:0044042	BP	glucan metabolic process	6	0.0082	0.39
GO:0006066	BP	alcohol metabolic process	8	0.0092	0.39
GO:0044264	BP	cellular polysaccharide metabolic process	6	0.012	0.44
GO:0009311	BP	oligosaccharide metabolic process	5	0.015	0.44
GO:0005984	BP	disaccharide metabolic process	5	0.014	0.44
GO:0005985	BP	sucrose metabolic process	5	0.013	0.44
GO:0016137	BP	glycoside metabolic process	5	0.015	0.44
GO:0005982	BP	starch metabolic process	5	0.017	0.46
GO:0019318	BP	hexose metabolic process	6	0.018	0.47
GO:0016051	BP	carbohydrate biosynthetic process	6	0.029	0.68
GO:0044271	BP	cellular nitrogen compound biosynthetic process	7	0.029	0.68
GO:0034637	BP	cellular carbohydrate biosynthetic process	5	0.034	0.76
GO:0006950	BP	response to stress	14	0.038	0.8
GO:0031982	CC	vesicle	92	0.0061	0.18
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	92	0.0059	0.18
GO:0031988	CC	membrane-bounded vesicle	92	0.0059	0.18
GO:0031410	CC	cytoplasmic vesicle	92	0.0061	0.18

Table S10: Significant GO terms for genes down-regulated in plants inoculated with isolate Sc2f compared to plants inoculated with isolate Sc2e. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0034641	BP	cellular nitrogen compound metabolic process	9	0.00054	**0.1**
GO:0019752	BP	carboxylic acid metabolic process	12	0.04	0.65
GO:0006520	BP	cellular amino acid metabolic process	9	0.04	0.65
GO:0044106	BP	cellular amine metabolic process	9	0.04	0.65
GO:0009308	BP	amine metabolic process	10	0.028	0.65
GO:0043436	BP	oxoacid metabolic process	12	0.04	0.65
GO:0005975	BP	carbohydrate metabolic process	8	0.022	0.65
GO:0042180	BP	cellular ketone metabolic process	12	0.045	0.65
GO:0006082	BP	organic acid metabolic process	12	0.04	0.65
GO:0005976	BP	polysaccharide metabolic process	5	0.0084	0.65
GO:0016053	BP	organic acid biosynthetic process	6	0.016	0.65
GO:0046394	BP	carboxylic acid biosynthetic process	6	0.016	0.65
GO:0006915	BP	apoptosis	5	0.044	0.65
GO:0020037	MF	heme binding	8	9.90E-06	**0.00078**
GO:0046906	MF	tetrapyrrole binding	8	1.50E-05	**0.00078**
GO:0005506	MF	iron ion binding	8	0.0015	**0.054**
GO:0016757	MF	transferase activity	7	0.0054	0.14
GO:0016740	MF	transferase activity	26	0.029	0.62
GO:0004175	MF	endopeptidase activity	5	0.082	1
GO:0031982	CC	vesicle	80	7.50E-09	**1.70E-07**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	80	7.10E-09	**1.70E-07**
GO:0031988	CC	membrane-bounded vesicle	80	7.10E-09	**1.70E-07**
GO:0031410	CC	cytoplasmic vesicle	80	7.50E-09	**1.70E-07**
GO:0005886	CC	plasma membrane	6	0.036	0.65

Table S11: Significant GO terms for genes up-regulated in plants inoculated with isolate Sc2b compared to plants inoculated with isolate Sc2a. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0015671	BP	oxygen transport	5	1.30E-09	**5.60E-07**
GO:0015669	BP	gas transport	5	4.50E-09	**9.70E-07**
GO:0034641	BP	cellular nitrogen compound metabolic process	17	4.40E-05	**0.0064**
GO:0006629	BP	lipid metabolic process	16	0.0011	0.11
GO:0006631	BP	fatty acid metabolic process	8	0.0016	0.14
GO:0032787	BP	monocarboxylic acid metabolic process	13	0.0047	0.21
GO:0043436	BP	oxoacid metabolic process	27	0.0074	0.21
GO:0016053	BP	organic acid biosynthetic process	12	0.0038	0.21
GO:0046394	BP	carboxylic acid biosynthetic process	12	0.0038	0.21
GO:0006633	BP	fatty acid biosynthetic process	6	0.0048	0.21
GO:0016310	BP	phosphorylation	23	0.0055	0.21
GO:0044255	BP	cellular lipid metabolic process	13	0.005	0.21
GO:0006082	BP	organic acid metabolic process	27	0.0074	0.21
GO:0019752	BP	carboxylic acid metabolic process	27	0.0074	0.21
GO:0007017	BP	microtubule-based process	5	0.0071	0.21
GO:0042180	BP	cellular ketone metabolic process	27	0.0093	0.25
GO:0005975	BP	carbohydrate metabolic process	15	0.018	0.41
GO:0006796	BP	phosphate metabolic process	23	0.018	0.41
GO:0006793	BP	phosphorus metabolic process	23	0.018	0.41
GO:0020037	MF	heme binding	16	4.30E-09	**9.30E-07**
GO:0046906	MF	tetrapyrrole binding	16	9.70E-09	**1.00E-06**
GO:0005506	MF	iron ion binding	20	3.60E-07	**2.60E-05**
GO:0019825	MF	oxygen binding	5	5.90E-07	**3.20E-05**
GO:0016491	MF	oxidoreductase activity	28	0.00032	**0.014**
GO:0016705	MF	oxidoreductase activity	5	0.0013	**0.047**
GO:0046914	MF	transition metal ion binding	36	0.0023	**0.07**
GO:0016798	MF	hydrolase activity	9	0.0039	**0.1**
GO:0004553	MF	compounds	8	0.0046	0.11
GO:0046872	MF	metal ion binding	43	0.0096	0.21
GO:0043169	MF	cation binding	44	0.012	0.21
GO:0043167	MF	ion binding	44	0.012	0.21
GO:0031982	CC	vesicle	152	3.60E-09	**1.30E-07**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	152	3.40E-09	**1.30E-07**
GO:0031988	CC	membrane-bounded vesicle	152	3.40E-09	**1.30E-07**
GO:0031410	CC	cytoplasmic vesicle	152	3.60E-09	**1.30E-07**
GO:0015630	CC	microtubule cytoskeleton	5	0.0016	**0.045**
GO:0005576	CC	extracellular region	7	0.0026	**0.06**
GO:0005856	CC	cytoskeleton	6	0.037	0.74
GO:0044430	CC	cytoskeletal part	5	0.045	0.79

Table S12: Significant GO terms for genes up-regulated in plants inoculated with isolate Sc2e compared to plants inoculated with isolate Sc2a. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0016998	BP	cell wall macromolecule catabolic process	6	3.50E-07	**0.00012**
GO:0044036	BP	cell wall macromolecule metabolic process	7	4.80E-07	**0.00012**
GO:0006022	BP	aminoglycan metabolic process	5	1.10E-05	**0.0015**
GO:0034641	BP	cellular nitrogen compound metabolic process	16	1.30E-05	**0.0015**
GO:0005975	BP	carbohydrate metabolic process	17	0.00045	**0.043**
GO:0005976	BP	polysaccharide metabolic process	9	0.00078	**0.062**
GO:0006950	BP	response to stress	19	0.0011	**0.072**
GO:0050896	BP	response to stimulus	27	0.0012	**0.072**
GO:0044262	BP	cellular carbohydrate metabolic process	11	0.0023	0.12
GO:0016052	BP	carbohydrate catabolic process	7	0.0034	0.16
GO:0008652	BP	cellular amino acid biosynthetic process	7	0.0037	0.16
GO:0009309	BP	amine biosynthetic process	7	0.006	0.24
GO:0009308	BP	amine metabolic process	18	0.0089	0.32
GO:0044271	BP	cellular nitrogen compound biosynthetic process	8	0.012	0.4
GO:0009790	BP	embryonic development	5	0.013	0.4
GO:0009066	BP	aspartate family amino acid metabolic process	5	0.015	0.45
GO:0051707	BP	response to other organism	5	0.018	0.46
GO:0016053	BP	organic acid biosynthetic process	9	0.018	0.46
GO:0046394	BP	carboxylic acid biosynthetic process	9	0.018	0.46
GO:0009607	BP	response to biotic stimulus	5	0.026	0.61
GO:0009628	BP	response to abiotic stimulus	8	0.032	0.7
GO:0006066	BP	alcohol metabolic process	7	0.031	0.7
GO:0006575	BP	cellular amino acid derivative metabolic process	7	0.036	0.74
GO:0006519	BP	cellular amino acid and derivative metabolic process	17	0.038	0.76
GO:0004197	MF	cysteine-type endopeptidase activity	5	1.70E-05	**0.0036**
GO:0004568	MF	chitinase activity	5	9.30E-05	**0.0099**
GO:0008234	MF	cysteine-type peptidase activity	5	0.00039	**0.028**
GO:0004553	MF	compounds	8	0.0013	**0.069**
GO:0016798	MF	hydrolase activity, acting on glycosyl bonds	8	0.0037	0.14
GO:0016491	MF	oxidoreductase activity	21	0.0039	0.14
GO:0003824	MF	catalytic activity	98	0.028	0.85
GO:0020037	MF	heme binding	5	0.046	0.93
GO:0016787	MF	hydrolase activity	34	0.036	0.93
GO:0005506	MF	iron ion binding	8	0.056	0.93
GO:0050662	MF	coenzyme binding	6	0.05	0.93
GO:0031982	CC	vesicle	100	0.0011	**0.052**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	100	0.0011	**0.052**
GO:0031410	CC	cytoplasmic vesicle	100	0.0011	**0.052**

Table S13: Significant GO terms for genes up-regulated in plants inoculated with isolate Sc2f compared to plants inoculated with isolate Sc2b. Column headings follow Table 1.

Probe sets	Function
a)	
Os.26226.1.S1_at	ATP binding protein
b)	
Os.8655.1.S1_at	Thionin precursor
Os.54146.1.S1_at	Nitrate reductase

Table S14: a) Probe sets that were up-regulated in plants inoculated with isolate S3c and down-regulated in plants inoculated with isolates S3a and S3b compared to the plants inoculated with the parental isolate S3. b) Probe sets that were down-regulated in plants inoculated with S3c and up-regulated in plants inoculated with isolates S3a and S3b compared to the plants inoculated with the parental isolate S3.

GO	Type	Functions	Count	p-value	FDR
GO:0050896	BP	response to stimulus	15	0.00048	**0.045**
GO:0032787	BP	monocarboxylic acid metabolic process	7	0.002	**0.065**
GO:0006950	BP	response to stress	10	0.0019	**0.065**
GO:0020037	MF	heme binding	5	0.00099	**0.055**
GO:0046906	MF	tetrapyrrole binding	5	0.0013	**0.055**
GO:0005506	MF	iron ion binding	6	0.0054	0.15
GO:0016491	MF	oxidoreductase activity	9	0.026	0.56
GO:0031982	CC	vesicle	46	0.0012	**0.019**
GO:0016023	CC	cytoplasmic membrane	46	0.0012	**0.019**
GO:0031410	CC	cytoplasmic vesicle	46	0.0012	**0.019**
GO:0031988	CC	membrane-bounded vesicle	46	0.0012	**0.019**

Table S15: Significant GO terms for genes down-regulated in plants inoculated with isolate S3a compared to plants inoculated with parental isolate S3. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0006090	BP	pyruvate metabolic process	6	1.30E-05	**0.0021**
GO:0034637	BP	cellular carbohydrate biosynthetic process	6	5.90E-05	**0.0046**
GO:0032787	BP	monocarboxylic acid metabolic process	8	0.00026	**0.01**
GO:0016051	BP	carbohydrate biosynthetic process	6	0.00023	**0.01**
GO:0044262	BP	cellular carbohydrate metabolic process	6	0.003	**0.078**
GO:0008610	BP	lipid biosynthetic process	5	0.0029	**0.078**
GO:0006725	BP	cellular aromatic compound metabolic process	6	0.0037	**0.082**
GO:0034641	BP	cellular nitrogen compound metabolic process	6	0.005	**0.087**
GO:0050896	BP	response to stimulus	12	0.0051	**0.087**
GO:0006091	BP	generation of precursor metabolites and energy	5	0.0077	0.12
GO:0005975	BP	carbohydrate metabolic process	7	0.0097	0.14
GO:0019752	BP	carboxylic acid metabolic process	10	0.02	0.22
GO:0043436	BP	oxoacid metabolic process	10	0.02	0.22
GO:0006082	BP	organic acid metabolic process	10	0.02	0.22
GO:0042180	BP	cellular ketone metabolic process	10	0.022	0.23
GO:0042221	BP	response to chemical stimulus	6	0.024	0.23
GO:0006629	BP	lipid metabolic process	5	0.04	0.37
GO:0006810	BP	transport	10	0.05	0.41
GO:0051234	BP	establishment of localization	10	0.05	0.41
GO:0020037	MF	heme binding	8	6.60E-07	**5.60E-05**
GO:0046906	MF	tetrapyrrole binding	8	1.00E-06	**5.60E-05**
GO:0005506	MF	iron ion binding	9	2.20E-05	**0.00081**
GO:0016491	MF	oxidoreductase activity	14	5.80E-05	**0.0016**
GO:0043167	MF	ion binding	19	0.002	**0.037**
GO:0043169	MF	cation binding	19	0.002	**0.037**
GO:0046872	MF	metal ion binding	18	0.0029	**0.045**
GO:0003824	MF	catalytic activity	40	0.025	0.35
GO:0046914	MF	transition metal ion binding	11	0.047	0.43
GO:0001883	MF	purine nucleoside binding	8	0.047	0.43
GO:0001882	MF	nucleoside binding	8	0.047	0.43
GO:0030554	MF	adenyl nucleotide binding	8	0.047	0.43
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	45	0.0005	**0.0092**
GO:0031410	CC	cytoplasmic vesicle	45	0.00051	**0.0092**
GO:0031982	CC	vesicle	45	0.00051	**0.0092**
GO:0031988	CC	membrane-bounded vesicle	45	0.0005	**0.0092**

Table S16: Significant GO terms for genes down-regulated in plants inoculated with isolate S3b compared to plants inoculated with parental isolate S3. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0016101	BP	diterpenoid metabolic process	7	7.60E-07	**0.00015**
GO:0006950	BP	response to stress	26	1.00E-06	**0.00015**
GO:0050896	BP	response to stimulus	36	7.80E-07	**0.00015**
GO:0009607	BP	response to biotic stimulus	10	8.40E-06	**0.00081**
GO:0006721	BP	terpenoid metabolic process	7	1.10E-05	**0.00081**
GO:0019748	BP	secondary metabolic process	12	1.10E-05	**0.00081**
GO:0006720	BP	isoprenoid metabolic process	7	6.40E-05	**0.004**
GO:0051704	BP	multi-organism process	10	8.70E-05	**0.0047**
GO:0044036	BP	cell wall macromolecule metabolic process	5	0.00012	**0.0056**
GO:0051707	BP	response to other organism	8	0.00016	**0.0065**
GO:0006952	BP	defense response	10	0.00016	**0.0065**
GO:0009698	BP	phenylpropanoid metabolic process	5	0.003	**0.1**
GO:0006629	BP	lipid metabolic process	13	0.0031	**0.1**
GO:0006575	BP	cellular amino acid derivative metabolic process	9	0.0043	0.13
GO:0034641	BP	cellular nitrogen compound metabolic process	11	0.0053	0.15
GO:0006118	BP	electron transport	9	0.0067	0.18
GO:0005975	BP	carbohydrate metabolic process	14	0.0077	0.2
GO:0005976	BP	polysaccharide metabolic process	7	0.011	0.27
GO:0044255	BP	cellular lipid metabolic process	10	0.019	0.42
GO:0042221	BP	response to chemical stimulus	12	0.024	0.52
GO:0009416	BP	response to light stimulus	5	0.03	0.63
GO:0009314	BP	response to radiation	5	0.034	0.67
GO:0032787	BP	monocarboxylic acid metabolic process	9	0.041	0.78
GO:0016052	BP	carbohydrate catabolic process	5	0.045	0.81
GO:0020037	MF	heme binding	19	2.80E-13	**5.20E-11**
GO:0046906	MF	tetrapyrrole binding	19	7.70E-13	**7.20E-11**
GO:0005506	MF	iron ion binding	22	5.40E-10	**3.40E-08**
GO:0043169	MF	cation binding	52	1.80E-06	**6.90E-05**
GO:0043167	MF	ion binding	52	1.80E-06	**6.90E-05**
GO:0046872	MF	metal ion binding	44	0.00015	**0.0048**
GO:0016491	MF	oxidoreductase activity	25	0.00018	**0.0049**
GO:0030246	MF	carbohydrate binding	6	0.0008	**0.019**
GO:0046914	MF	transition metal ion binding	31	0.0021	**0.043**
GO:0016798	MF	hydrolase activity	7	0.013	0.25
GO:0004553	MF	hydrolase activity	6	0.019	0.33
GO:0016758	MF	transferase activity, transferring hexosyl groups	7	0.032	0.5
GO:0031982	CC	vesicle	134	5.50E-10	**1.90E-08**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	134	5.10E-10	**1.90E-08**
GO:0031988	CC	membrane-bounded vesicle	134	5.10E-10	**1.90E-08**
GO:0031410	CC	cytoplasmic vesicle	134	5.50E-10	**1.90E-08**
GO:0005576	CC	extracellular region	5	0.017	0.47

Table S17: Significant GO terms for genes down-regulated in plants inoculated with isolate S3c compared to plants inoculated with parental isolate S3. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0006950	BP	response to stress	20	0.00027	**0.092**
GO:0009607	BP	response to biotic stimulus	7	0.0012	0.15
GO:0050896	BP	response to stimulus	26	0.0014	0.15
GO:0051707	BP	response to other organism	6	0.0035	0.3
GO:0070887	BP	cellular response to chemical stimulus	5	0.0093	0.52
GO:0006952	BP	defense response	7	0.0079	0.52
GO:0006118	BP	electron transport	8	0.015	0.71
GO:0051704	BP	multi-organism process	6	0.019	0.8
GO:0020037	MF	heme binding	14	9.50E-09	**1.90E-06**
GO:0046906	MF	tetrapyrrole binding	14	2.00E-08	**1.90E-06**
GO:0005506	MF	iron ion binding	15	1.50E-05	**0.00099**
GO:0004866	MF	endopeptidase inhibitor activity	5	5.50E-05	**0.0022**
GO:0030414	MF	peptidase inhibitor activity	5	5.50E-05	**0.0022**
GO:0004857	MF	enzyme inhibitor activity	5	0.00032	**0.011**
GO:0043169	MF	cation binding	42	0.00044	**0.011**
GO:0043167	MF	ion binding	42	0.00044	**0.011**
GO:0016758	MF	transferase activity, transferring hexosyl groups	9	0.0026	**0.057**
GO:0046872	MF	metal ion binding	37	0.0035	**0.07**
GO:0016491	MF	oxidoreductase activity	20	0.0051	**0.091**
GO:0016757	MF	transferase activity, transferring glycosyl groups	10	0.0061	**0.1**
GO:0030528	MF	transcription regulator activity	15	0.0075	0.11
GO:0030234	MF	enzyme regulator activity	5	0.0071	0.11
GO:0046914	MF	transition metal ion binding	27	0.011	0.14
GO:0004497	MF	monooxygenase activity	5	0.034	0.42
GO:0003700	MF	transcription factor activity	10	0.047	0.55
GO:0004553	MF	hydrolase activity	5	0.05	0.55
GO:0031982	CC	vesicle	121	4.00E-08	**1.10E-06**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	121	3.80E-08	**1.10E-06**
GO:0031988	CC	membrane-bounded vesicle	121	3.80E-08	**1.10E-06**
GO:0031410	CC	cytoplasmic vesicle	121	4.00E-08	**1.10E-06**
GO:0005576	CC	extracellular region	7	0.00061	**0.013**

Table S18: Significant GO terms for genes down-regulated in plants inoculated with isolate S3c compared to plants inoculated with isolate S3a. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0044036	BP	cell wall macromolecule metabolic process	8	4.90E-08	**2.30E-05**
GO:0016998	BP	cell wall macromolecule catabolic process	6	6.00E-07	**0.00014**
GO:0006022	BP	aminoglycan metabolic process	5	1.60E-05	**0.0025**
GO:0034641	BP	cellular nitrogen compound metabolic process	16	3.70E-05	**0.0042**
GO:0006118	BP	electron transport	13	8.90E-05	**0.0082**
GO:0016101	BP	diterpenoid metabolic process	5	0.00023	**0.017**
GO:0006950	BP	response to stress	21	0.00056	**0.037**
GO:0006721	BP	terpenoid metabolic process	5	0.0013	**0.078**
GO:0050896	BP	response to stimulus	28	0.0021	**0.096**
GO:0019748	BP	secondary metabolic process	9	0.0019	**0.096**
GO:0018130	BP	heterocycle biosynthetic process	5	0.0024	**0.1**
GO:0006720	BP	isoprenoid metabolic process	5	0.0045	0.17
GO:0009607	BP	response to biotic stimulus	6	0.01	0.36
GO:0006952	BP	defense response	7	0.015	0.48
GO:0006725	BP	cellular aromatic compound metabolic process	10	0.016	0.48
GO:0005976	BP	polysaccharide metabolic process	7	0.017	0.49
GO:0006575	BP	cellular amino acid derivative metabolic process	8	0.021	0.56
GO:0006979	BP	response to oxidative stress	5	0.024	0.61
GO:0051707	BP	response to other organism	5	0.025	0.62
GO:0051704	BP	multi-organism process	6	0.033	0.67
GO:0005506	MF	iron ion binding	22	2.30E-09	**5.40E-07**
GO:0020037	MF	heme binding	15	6.40E-09	**7.40E-07**
GO:0046906	MF	tetrapyrrole binding	15	1.40E-08	**1.10E-06**
GO:0043169	MF	cation binding	58	2.20E-07	**1.00E-05**
GO:0043167	MF	ion binding	58	2.20E-07	**1.00E-05**
GO:0046872	MF	metal ion binding	52	5.20E-06	**0.0002**
GO:0004568	MF	chitinase activity	5	0.00014	**0.0045**
GO:0046914	MF	transition metal ion binding	37	0.00016	**0.0045**
GO:0030246	MF	carbohydrate binding	7	0.00019	**0.0049**
GO:0004553	MF	hydrolase activity	9	0.00054	**0.012**
GO:0005507	MF	copper ion binding	5	0.0013	**0.028**
GO:0016798	MF	hydrolase activity, acting on glycosyl bonds	9	0.0018	**0.034**
GO:0005509	MF	calcium ion binding	9	0.0028	**0.05**
GO:0050660	MF	FAD binding	5	0.003	**0.05**
GO:0016757	MF	transferase activity, transferring glycosyl groups	11	0.0054	**0.083**
GO:0016758	MF	transferase activity, transferring hexosyl groups	9	0.0061	**0.087**
GO:0016491	MF	oxidoreductase activity	21	0.01	0.14
GO:0004175	MF	endopeptidase activity	11	0.011	0.14
GO:0003700	MF	transcription factor activity	12	0.023	0.28
GO:0031982	CC	vesicle	129	3.70E-07	**1.40E-05**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	129	3.50E-07	**1.40E-05**
GO:0031988	CC	membrane-bounded vesicle	129	3.50E-07	**1.40E-05**
GO:0031410	CC	cytoplasmic vesicle	129	3.70E-07	**1.40E-05**
GO:0005576	CC	extracellular region	6	0.006	**0.18**

Table S19: Significant GO terms for genes down-regulated in plants inoculated with isolate S3c compared to plants inoculated with isolate S3b. Column headings follow Table 1.

Probe sets	Function
a) Down regulated	
OsAffx.22999.1.S1_at	Expressed protein
Os.43929.1.S1_s_at	Expressed protein
Os.8823.1.S1_at	Expressed protein
Os.11914.1.S1_at	12-oxophytodienoate reductase 2
Os.14210.1.S1_at	Expressed protein
OsAffx.2364.2.S1_at	Hypothetical protein
Os.51647.1.S1_at	Expressed protein
OsAffx.2017.1.S1_at	Unknown protein
Os.55868.1.S1_at	Endo-1,4-beta-xylanase
OsAffx.28660.1.A1_at	Conserved hypothetical protein
b) Up regulated	
OsAffx.12905.1.S1_at	Hypothetical protein

Table S20: a) Probe sets that were commonly down-regulated in plants inoculated with isolates S4a, S4b or S4c compared to the plants inoculated with parental isolate S4. b) Probe set that was commonly up-regulated in plants inoculated with isolates S4a, S4b or S4c compared to the plants inoculated with parental isolate S4.

GO	Type	Functions	Count	p-value	FDR
GO:0016101	BP	diterpenoid metabolic process	5	5.80E-05	**0.018**
GO:0006721	BP	terpenoid metabolic process	5	0.00037	**0.058**
GO:0006720	BP	isoprenoid metabolic process	5	0.0013	**0.1**
GO:0019748	BP	secondary metabolic process	8	0.0011	**0.1**
GO:0006629	BP	lipid metabolic process	11	0.0046	0.29
GO:0016491	MF	oxidoreductase activity	21	0.00037	**0.056**
GO:0031982	CC	vesicle	88	0.00031	**0.0079**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	88	0.0003	**0.0079**
GO:0031988	CC	membrane-bounded vesicle	88	0.0003	**0.0079**
GO:0031410	CC	cytoplasmic vesicle	88	0.00031	**0.0079**

Table S21: Significant GO terms for genes altered in their transcription in plants inoculated with isolate S4b compared to plants inoculated with parental isolate S4. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0016998	BP	cell wall macromolecule catabolic process	8	1.10E-10	**3.70E-08**
GO:0044036	BP	cell wall macromolecule metabolic process	8	8.90E-09	**1.50E-06**
GO:0016101	BP	diterpenoid metabolic process	8	1.40E-08	**1.60E-06**
GO:0006032	BP	chitin catabolic process	5	4.70E-08	**3.90E-06**
GO:0006030	BP	chitin metabolic process	5	2.20E-07	**1.20E-05**
GO:0006026	BP	aminoglycan catabolic process	5	2.20E-07	**1.20E-05**
GO:0006721	BP	terpenoid metabolic process	8	3.20E-07	**1.50E-05**
GO:0006720	BP	isoprenoid metabolic process	8	2.70E-06	**0.00011**
GO:0006022	BP	aminoglycan metabolic process	5	5.50E-06	**0.0002**
GO:0006629	BP	lipid metabolic process	16	2.40E-05	**0.00074**
GO:0009685	BP	gibberellin metabolic process	5	2.30E-05	**0.00074**
GO:0000272	BP	polysaccharide catabolic process	5	0.00012	**0.0034**
GO:0019748	BP	secondary metabolic process	9	0.0004	**0.0095**
GO:0051707	BP	response to other organism	7	0.0004	**0.0095**
GO:0009607	BP	response to biotic stimulus	7	0.00068	**0.015**
GO:0051704	BP	multi-organism process	8	0.00073	**0.015**
GO:0044255	BP	cellular lipid metabolic process	12	0.00084	**0.016**
GO:0050896	BP	response to stimulus	24	0.0017	**0.032**
GO:0034641	BP	cellular nitrogen compound metabolic process	10	0.0056	**0.098**
GO:0006950	BP	response to stress	15	0.0084	0.14
GO:0005976	BP	polysaccharide metabolic process	6	0.019	0.3
GO:0009308	BP	amine metabolic process	15	0.023	0.31
GO:0016053	BP	organic acid biosynthetic process	8	0.022	0.31
GO:0046394	BP	carboxylic acid biosynthetic process	8	0.022	0.31
GO:0016052	BP	carbohydrate catabolic process	5	0.027	0.36
GO:0008610	BP	lipid biosynthetic process	6	0.031	0.4
GO:0019752	BP	carboxylic acid metabolic process	18	0.04	0.46
GO:0043436	BP	oxoacid metabolic process	18	0.04	0.46
GO:0006082	BP	organic acid metabolic process	18	0.04	0.46
GO:0032787	BP	monocarboxylic acid metabolic process	8	0.047	0.5
GO:0042180	BP	cellular ketone metabolic process	18	0.046	0.5
GO:0004568	MF	chitinase activity	5	4.90E-05	**0.0082**
GO:0016491	MF	oxidoreductase activity	23	0.00014	**0.012**
GO:0004497	MF	monooxygenase activity	7	0.0014	**0.056**
GO:0005506	MF	iron ion binding	11	0.0012	**0.056**
GO:0030246	MF	carbohydrate binding	5	0.0026	**0.086**
GO:0020037	MF	heme binding	6	0.0076	0.21
GO:0046906	MF	tetrapyrrole binding	6	0.0099	0.23
GO:0004553	MF	hydrolase activity	5	0.036	0.75
GO:0031982	CC	vesicle	118	2.70E-09	**6.70E-08**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	118	2.50E-09	**6.70E-08**
GO:0031988	CC	membrane-bounded vesicle	118	2.50E-09	**6.70E-08**
GO:0031410	CC	cytoplasmic vesicle	118	2.70E-09	**6.70E-08**

Table S22: Significant GO terms for genes down-regulated in plants inoculated with isolate S4c compared to plants inoculated with isolate S4b. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0005976	BP	polysaccharide metabolic process	7	0.00067	**0.078**
GO:0050896	BP	response to stimulus	20	0.00045	**0.078**
GO:0034641	BP	cellular nitrogen compound metabolic process	9	0.0015	0.11
GO:0051707	BP	response to other organism	5	0.0023	0.13
GO:0009607	BP	response to biotic stimulus	5	0.0033	0.15
GO:0042221	BP	response to chemical stimulus	10	0.0046	0.18
GO:0006950	BP	response to stress	12	0.0057	0.19
GO:0051704	BP	multi-organism process	5	0.01	0.3
GO:0009308	BP	amine metabolic process	12	0.014	0.36
GO:0033036	BP	macromolecule localization	5	0.03	0.69
GO:0006575	BP	cellular amino acid derivative metabolic process	5	0.038	0.8
GO:0005975	BP	carbohydrate metabolic process	8	0.046	0.86
GO:0006118	BP	electron transport	5	0.048	0.86
GO:0006519	BP	cellular amino acid process	11	0.054	0.89
GO:0004568	MF	chitinase activity	5	8.20E-06	**0.0013**
GO:0016491	MF	oxidoreductase activity	17	0.00045	**0.035**
GO:0005506	MF	iron ion binding	9	0.00098	**0.051**
GO:0004553	MF	hydrolase activity	6	0.0017	**0.067**
GO:0016798	MF	hydrolase activity	6	0.004	0.12
GO:0016758	MF	transferase activity, transferring hexosyl groups	6	0.0096	0.25
GO:0016757	MF	transferase activity, transferring glycosyl groups	7	0.011	0.26
GO:0022804	MF	active transmembrane transporter activity	6	0.02	0.38
GO:0003824	MF	catalytic activity	62	0.03	0.51
GO:0016787	MF	hydrolase activity	21	0.065	0.83
GO:0048037	MF	cofactor binding	5	0.059	0.83
GO:0030554	MF	adenyl nucleotide binding	11	0.074	0.83
GO:0001883	MF	purine nucleoside binding	11	0.074	0.83
GO:0001882	MF	nucleoside binding	11	0.074	0.83
GO:0031982	CC	vesicle	65	0.0014	**0.036**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	65	0.0014	**0.036**
GO:0031988	CC	membrane-bounded vesicle	65	0.0014	**0.036**
GO:0031410	CC	cytoplasmic vesicle	65	0.0014	**0.036**
GO:0031224	CC	intrinsic to membrane	10	0.021	0.42
GO:0031975	CC	envelope	5	0.029	0.42
GO:0031967	CC	organelle envelope	5	0.028	0.42

Table S23: Significant GO terms for genes down-regulated in plants inoculated with isolate S4c compared to plants inoculated with isolate S4a. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0006950	BP	response to stress	28	0.00015	**0.072**
GO:0043169	MF	cation binding	61	0.00026	**0.033**
GO:0043167	MF	ion binding	61	0.00027	**0.033**
GO:0016705	MF	oxidoreductase activity	6	0.00049	**0.04**
GO:0046872	MF	metal ion binding	56	0.0012	**0.073**
GO:0016491	MF	oxidoreductase activity	29	0.0032	0.16
GO:0004553	MF	hydrolase activity	8	0.015	0.52
GO:0016798	MF	hydrolase activity	9	0.014	0.52
GO:0016758	MF	transferase activity	9	0.04	0.9
GO:0020037	MF	heme binding	7	0.03	0.9
GO:0046906	MF	tetrapyrrole binding	7	0.039	0.9
GO:0046914	MF	transition metal ion binding	36	0.039	0.9
GO:0030528	MF	transcription regulator activity	18	0.044	0.91
GO:0031982	CC	vesicle	151	0.00011	**0.0044**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	151	0.00011	**0.0044**
GO:0031988	CC	membrane-bounded vesicle	151	0.00011	**0.0044**
GO:0031410	CC	cytoplasmic vesicle	151	0.00011	**0.0044**

Table S24: Significant GO terms for genes altered in their transcription in plants inoculated with parental isolate S4 compared to plants inoculated with parental isolate S3. Column headings follow Table 1.

GO	Type	Functions	Count	p-value	FDR
GO:0050896	BP	response to stimulus	56	3.20E-10	**2.50E-07**
GO:0006950	BP	response to stress	36	8.00E-08	**3.10E-05**
GO:0016998	BP	cell wall macromolecule catabolic process	7	2.00E-07	**5.20E-05**
GO:0006032	BP	chitin catabolic process	5	7.30E-07	**0.00014**
GO:0006026	BP	aminoglycan catabolic process	5	3.40E-06	**0.00044**
GO:0006030	BP	chitin metabolic process	5	3.40E-06	**0.00044**
GO:0032787	BP	monocarboxylic acid metabolic process	21	1.00E-05	**0.00064**
GO:0051704	BP	multi-organism process	14	8.30E-06	**0.00064**
GO:0009607	BP	response to biotic stimulus	12	1.10E-05	**0.00064**
GO:0044036	BP	cell wall macromolecule metabolic process	7	7.60E-06	**0.00064**
GO:0006629	BP	lipid metabolic process	23	1.00E-05	**0.00064**
GO:0019252	BP	starch biosynthetic process	5	8.10E-06	**0.00064**
GO:0019748	BP	secondary metabolic process	15	8.90E-06	**0.00064**
GO:0005975	BP	carbohydrate metabolic process	26	1.60E-05	**0.00089**
GO:0051707	BP	response to other organism	11	2.40E-05	**0.0012**
GO:0005976	BP	polysaccharide metabolic process	14	2.40E-05	**0.0012**
GO:0006090	BP	pyruvate metabolic process	10	7.10E-05	**0.0033**
GO:0006022	BP	aminoglycan metabolic process	5	7.80E-05	**0.0034**
GO:0016101	BP	diterpenoid metabolic process	6	0.00012	**0.0048**
GO:0006720	BP	isoprenoid metabolic process	8	0.00013	**0.0052**
GO:0006721	BP	terpenoid metabolic process	7	0.00014	**0.0053**
GO:0019438	BP	aromatic compound biosynthetic process	9	0.00035	**0.012**
GO:0044262	BP	cellular carbohydrate metabolic process	16	0.00043	**0.015**
GO:0006118	BP	electron transport	14	0.00064	**0.021**
GO:0044255	BP	cellular lipid metabolic process	17	0.00081	**0.025**
GO:0042398	BP	cellular amino acid derivative biosynthetic process	9	0.00094	**0.028**
GO:0008610	BP	lipid biosynthetic process	12	0.00098	**0.028**
GO:0006575	BP	cellular amino acid derivative metabolic process	13	0.001	**0.028**
GO:0034641	BP	cellular nitrogen compound metabolic process	16	0.0013	**0.036**
GO:0009250	BP	glucan biosynthetic process	5	0.0014	**0.037**
GO:0000272	BP	polysaccharide catabolic process	5	0.0015	**0.038**
GO:0006073	BP	cellular glucan metabolic process	9	0.0018	**0.042**
GO:0044042	BP	glucan metabolic process	9	0.0018	**0.042**
GO:0009628	BP	response to abiotic stimulus	14	0.0018	**0.042**
GO:0018130	BP	heterocycle biosynthetic process	6	0.0019	**0.042**
GO:0005982	BP	starch metabolic process	8	0.0024	**0.053**
GO:0016052	BP	carbohydrate catabolic process	9	0.0029	**0.062**
GO:0044264	BP	cellular polysaccharide metabolic process	9	0.003	**0.062**
GO:0033692	BP	cellular polysaccharide biosynthetic process	5	0.0035	**0.069**
GO:0042221	BP	response to chemical stimulus	19	0.0035	**0.069**
GO:0009698	BP	phenylpropanoid metabolic process	6	0.0037	**0.071**
GO:0006725	BP	cellular aromatic compound metabolic process	14	0.0046	**0.086**
GO:0000271	BP	polysaccharide biosynthetic process	5	0.0065	0.12
GO:0016053	BP	organic acid biosynthetic process	13	0.0076	0.13
GO:0046394	BP	carboxylic acid biosynthetic process	13	0.0076	0.13
GO:0034637	BP	cellular carbohydrate biosynthetic process	8	0.0077	0.13
GO:0009060	BP	aerobic respiration	5	0.01	0.17
GO:0006952	BP	defense response	9	0.01	0.17
GO:0016051	BP	carbohydrate biosynthetic process	9	0.011	0.18
GO:0015980	BP	energy derivation by oxidation of organic compounds	6	0.013	0.2
GO:0043648	BP	dicarboxylic acid metabolic process	5	0.014	0.21

GO:0005996	BP	monosaccharide metabolic process	9	0.014	0.21
GO:0019318	BP	hexose metabolic process	8	0.017	0.25
GO:0009416	BP	response to light stimulus	7	0.017	0.25
GO:0009314	BP	response to radiation	7	0.02	0.28
GO:0005985	BP	sucrose metabolic process	6	0.023	0.32
GO:0005984	BP	disaccharide metabolic process	6	0.025	0.34
GO:0015976	BP	carbon utilization	5	0.026	0.34
GO:0009311	BP	oligosaccharide metabolic process	6	0.026	0.34
GO:0016137	BP	glycoside metabolic process	6	0.026	0.34
GO:0043436	BP	oxoacid metabolic process	29	0.029	0.36
GO:0006082	BP	organic acid metabolic process	29	0.029	0.36
GO:0019752	BP	carboxylic acid metabolic process	29	0.029	0.36
GO:0045333	BP	cellular respiration	5	0.033	0.41
GO:0006006	BP	glucose metabolic process	6	0.035	0.42
GO:0042180	BP	cellular ketone metabolic process	29	0.036	0.43
GO:0006091	BP	generation of precursor metabolites and energy	10	0.037	0.43
GO:0006066	BP	alcohol metabolic process	9	0.041	0.47
GO:0006633	BP	fatty acid biosynthetic process	5	0.043	0.49
GO:0016491	MF	oxidoreductase activity	46	2.20E-09	**5.90E-07**
GO:0020037	MF	heme binding	14	2.30E-06	**0.00031**
GO:0046906	MF	tetrapyrrole binding	14	4.50E-06	**0.0004**
GO:0016705	MF	oxidoreductase activity	8	7.30E-06	**0.00049**
GO:0016762	MF	xyloglucan:xyloglucosyl transferase activity	5	3.20E-05	**0.0017**
GO:0004568	MF	chitinase activity	6	6.60E-05	**0.003**
GO:0005506	MF	iron ion binding	18	8.40E-05	**0.0032**
GO:0004497	MF	monooxygenase activity	11	0.00016	**0.0053**
GO:0016627	MF	oxidoreductase activity	6	0.00046	**0.014**
GO:0043167	MF	ion binding	58	0.0012	**0.033**
GO:0043169	MF	cation binding	57	0.0019	**0.046**
GO:0050662	MF	coenzyme binding	11	0.0026	**0.059**
GO:0004553	MF	hydrolase activity	9	0.005	**0.097**
GO:0016798	MF	hydrolase activity	10	0.005	**0.097**
GO:0030246	MF	carbohydrate binding	6	0.0062	0.11
GO:0048037	MF	cofactor binding	12	0.008	0.13
GO:0016830	MF	carbon-carbon lyase activity	5	0.016	0.26
GO:0046872	MF	metal ion binding	49	0.022	0.31
GO:0003824	MF	catalytic activity	145	0.022	0.31
GO:0016829	MF	lyase activity	10	0.03	0.4
GO:0016616	MF	oxidoreductase activity	5	0.038	0.48
GO:0016758	MF	transferase activity	9	0.041	0.49
GO:0030234	MF	enzyme regulator activity	5	0.042	0.49
GO:0031982	CC	vesicle	185	2.20E-10	**9.90E-09**
GO:0016023	CC	cytoplasmic membrane-bounded vesicle	185	2.10E-10	**9.90E-09**
GO:0031988	CC	membrane-bounded vesicle	185	2.10E-10	**9.90E-09**
GO:0031410	CC	cytoplasmic vesicle	185	2.20E-10	**9.90E-09**

Table S25: Significant GO terms for genes altered in their transcription in plants inoculated with parental isolate Sc2 compared to plants inoculated with parental isolate S3. Column headings follow Table 1.

Genes	Function
OsAM84	Putative multidrug resistance protein 1
OsAM189	Putative flavonoid 3',5'-hydroxylase
OsAM213	Hypothetical protein

Table S26: Gene commonly altered in their transcription in our experiment and in Guimil *et al.* experiment (Guimil *et al.* 2005). Gene in bold is down-regulated in the data from Gumil *et al.* (Guimil *et al.* 2005).

Generale discussion and perspectives

Genetic variation exists naturally among individuals coexisting in AMF populations (Koch *et al.* 2004). A recent study demonstrated that genetic exchange occurs between genetically different AMF isolates (Croll *et al.* 2009). Additionally, Angelard *et al.* (2010) have reported that segregation can occur in AMF during spore formation. Until now, little was known about genetic exchange and segregation and their consequences on AMF and their host plants. Nevertheless, segregation has been shown to alter AM specific plant gene transcription, to influence plant growth (Angelard *et al.* 2010) and to affect plant root colonization (Angelard *et al.* 2011).

The main goal of my thesis was to further investigate the effects of this genetic variation in AMF and in particular, to determine the impact of these two processes (genetic exchange and segregation) on the transcriptome of *Oryza sativa*, one of the most important crops worldwide.

Altogether, the results presented in my PhD thesis demonstrate that genetic variation in AMF, which can be enhanced by genetic exchange and segregation, influences their host and affects gene transcription.

The experiments presented in Chapter 1 show that genetic variation caused by genetic exchange has important consequences for their symbiotic effects on plants. Indeed, genetic exchange in AMF modified the transcription levels of some AM-specific plant genes. In addition, the time course of root colonization differed among plants inoculated with different crossed AMF lines.

In Chapter 2, I compared the transcriptome profile of rice inoculated with two genetically different AMF isolates with the transcriptome profile of uninoculated plants. This analysis showed that plants did not share the same AM symbiosis specific genes when they are inoculated with two different AMF isolates. In addition, comparison between plants inoculated with two different isolates allowed me to detect alterations in metabolic pathways that were not detected in the traditional mycorrhizal / non-mycorrhizal comparisons.

In Chapter 3, I clearly showed that segregation in AMF affects the transcription level of other plant genes than those described in Angelard *et al.* (2010). Most notably these were genes involved in plant defence or carbohydrate metabolism. Additionally, segregated lines coming from the same parental line led to different effects on plant transcriptome from each other. Moreover, I could observed that there was no additional differences in transcriptome profiles between plants inoculated with two parental lines than between plants inoculated with two segregated lines originating from the same parental line.

It appears that environmental or experimental conditions determine which AM-specific genes were identified. In particular, the timing of colonization at which the transcriptome analysis is made seems to be an important factor. In Chapter 1, I demonstrated that the time course of colonization depending on AMF genetic variation has an effect on plant gene expression. In accordance, Gutjahr *et al.* (2008) have shown that some plant genes are expressed only in the earlier stages of the colonization. A study made in 2003 which examined transcript profiles of *Medicago truncatula* inoculated between 10 and 38 days with the AMF *Glomus versiforme* using cDNA arrays (Liu *et al.* 2003) have shown that genes involved in plant defences are firstly up-regulated in the early stages of the colonization and then down-regulated. These results suggest that the profile of gene expression is not the same pattern at the beginning of colonization and when the colonization is well established. This is in accordance with Gutjahr *et al.* (2008) and Balestrini *et al.*(2006). In 2005, Guimil *et al.* established a list of 224 AM-specific genes that were only transcribed in rice in the presence of AMF (Guimil *et al.* 2005). In chapter 2, by comparing the transcriptome profiles of rice inoculated with different AMF isolates with the transcriptome profiles of uninoculated plants, I found that the list of AM-specific genes identified was different. However, it is important to note that in my study the array hybridization was performed using transcripts extracted from plant roots 12 weeks after inoculation, while Guimil *et al.* (2005) carried out their microarrays hybridization from roots harvested after 6 weeks of colonization by AMF. This difference probably explains the discrepancies.

AMF inter-specific genetic variability was already shown to largely influence the transcriptome of plants (Hohnjec *et al.* 2005; Grunwald *et al.* 2009). In chapter 2, I studied the effects of intra-specific genetic variability on the rice transcriptome. My results demonstrated, for the first time, that the intra-specific genetic variability influences the host transcriptome as much as the inter-specific genetic variability. Interestingly, among the AM-specific genes identified, some pathways are regularly altered such as those involved in stress, defence, carbohydrate metabolism or ion binding and transport. Remarkably, in concordance with my findings in Chapters 2 and 3, Liu *et al.* (2003) also described that genes involved in plant defences were down-regulated in later stages of colonization. In 2005, Hohnjec *et al.* (2005) analysed the differences in transcription profiles between *M. Truncatula* inoculated with *G. intraradices* or *Glomus mosseae* 28 days after inoculation. They found that 201 plant genes were coinduced by both of the fungi, but a further 453 plant genes were induced specifically by *G. mosseae* and 556 specifically induced by *G. intraradices*, indicating that inter-species

genetic variability have a direct effect on host transcriptome. In this study, genes belonging to pathways involved in cell wall, defence, nutrient and ion transport, stress and cytochrome P450 were commonly altered in plants inoculated with an AMF compared to uninoculated plants. Strikingly, these pathways are similar to those I found in the mycorrhizal/non-mycorrhizal comparison in Chapters 2 and 3.

In 2009, Grunwald *et al.* (2009) compared the transcriptome profile of *M. truncatula* that grew with the AMF species *Gigaspora rosea* and control plants, six weeks after inoculation. They contrasted their results with the results from Hohjnec *et al.* (Hohnjec *et al.* 2005). Very few genes were identified as being commonly regulated by all three AMF. It appeared that each AMF has specific effects on plant transcriptomes. Considering that these species are relatively distant from a phylogenetic point of view, it was somewhat surprising that each species differently influences the host. In chapter 3, unexpectedly we found that plant inoculated with different segregated lines also exhibited a large number of specifically transcribed genes in rice that were different even though they are closely related from each other.

Perspectives

In my thesis, I have largely studied the consequences of AMF genetic variability on the plant transcriptome. My results have led to a better understanding of AMF/plant symbiotic interactions at the gene expression level. However, several questions remain unsolved, especially those concerning the fungal partner. How could segregation and genetic exchange affect AMF themselves? Effects on fungal phenotypic traits have already been studied (Angelard *et al.* 2011), but how is fungal gene expression affected? Do segregation and genetic exchange greatly influence the AMF transcriptome?

Because of the difficulty to obtain clean DNA or RNA from these fungi, very few genes are known and studied. Recently, the transcriptome of *G. intraradices* was released (E Tisserant *et al.*, *in press*), and a microarray containing 22 294 predicted cDNAs will be available. This new tool will permit us to study the consequences of genetic exchange and segregation on the AMF transcriptome.

In chapter 2, I used a different approach to the classical inoculated versus uninoculated plants, to detect new genes that were specifically sensitive to the AM symbiosis. These genes were

not present in the previous set of AM-specific genes identified by Guimil *et al.* The characterization and identification of their role in symbiosis have to be determined precisely, as was performed by Gutjahr *et al.* for the previous sets of genes (2008).

Finally, AMF and segregation in AMF have been shown to increase rice growth (Angelard *et al.* 2010). Rice is one of the most important food crops on Earth. Consequently, these fungi could be very useful from an agricultural point of view. Indeed, with the ecological movement that tends to use less chemical fertilizers, AMF present the advantage to be natural and non-toxic for human or for ecosystems. Nevertheless, the effects of genetic exchange and segregation remain for the moment, too unpredictable and deeper investigations need to be conducted in order to use these modified isolates as natural fertilizer. My results demonstrated that the AM-induced plant transcriptome variations are almost unlimited, thus exploiting this natural AMF variation might be useful to develop AMF inocula that are more beneficial for crop growth.

Bibliography

Al-Karaki, G. N. (1998). "Benefit, cost and water-use efficiency of arbuscular mycorrhizal durum wheat grown under drought stress." Mycorrhiza **8**(1): 41-45.

Angelard, C. (2010). Effect of genetic alteration in arbuscular mycorrhizal fungi on plant growth and on their response to a change of host, Lausanne. **Ph.D.**

Angelard, C., A. Colard, et al. (2010). "Segregation in a mycorrhizal fungus alters rice growth and symbiosis-specific gene transcription." Curr Biol **20**(13): 1216-1221.

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

Bago, B., P. E. Pfeffer, et al. (2003). "Carbon export from arbuscular mycorrhizal roots involves the translocation of carbohydrate as well as lipid." Plant Physiol **131**(3): 1496-1507.

Balestrini, R. and L. Lanfranco (2006). "Fungal and plant gene expression in arbuscular mycorrhizal symbiosis." Mycorrhiza **16**(8): 509-524.

Becard, G. and J. A. Fortin (1988). "Early Events of Vesicular Arbuscular Mycorrhiza Formation on Ri T-DNA Transformed Roots." New Phytologist **108**(2): 211-218.

Bever, J. D. and M. Wang (2005). "Arbuscular mycorrhizal fungi: hyphal fusion and multigenomic structure." Nature **433**(7022): E3-4; discussion E4.

Bianciotto, V., C. Bandi, et al. (1996). "An obligately endosymbiotic mycorrhizal fungus itself harbors obligately intracellular bacteria." Appl Environ Microbiol **62**(8): 3005-3010.

Breuillin, F., J. Schramm, et al. (2010). "Phosphate systemically inhibits development of arbuscular mycorrhiza in Petunia hybrida and represses genes involved in mycorrhizal functioning." Plant Journal **64**(6): 1002-1017.

Brundrett, M. C. (2002). "Coevolution of roots and mycorrhizas of land plants." New Phytologist **154**(2): 275-304.

Burggraaf, A. J. P. and J. E. Beringer (1989). "Absence of Nuclear-DNA Synthesis in Vesicular Arbuscular Mycorrhizal Fungi During Invitro Development." New Phytologist **111**(1): 25-33.

Caicedo, A. L., S. H. Williamson, et al. (2007). "Genome-wide patterns of nucleotide polymorphism in domesticated rice." PLoS Genet **3**(9): 1745-1756.

Campos-Soriano, L., J. M. Garcia-Garrido, et al. (2010). "Activation of basal defense mechanisms of rice plants by Glomus intraradices does not affect the arbuscular mycorrhizal symbiosis." New Phytol **188**(2): 597-614.

Clapp, J. P., A. H. Fitter, et al. (1999). "Ribosomal small subunit sequence variation within spores of an arbuscular mycorrhizal fungus, Scutellospora sp." Molecular Ecology **8**(6): 915-921.

Corradi, N., D. Croll, et al. (2007). "Gene copy number polymorphisms in an arbuscular mycorrhizal fungal population." Appl Environ Microbiol **73**(1): 366-369.

Corradi, N., G. Kuhn, et al. (2004). "Monophyly of beta-tubulin and H+-ATPase gene variants in Glomus intraradices: consequences for molecular evolutionary studies of AM fungal genes." Fungal Genetics and Biology **41**(2): 262-273.

Corradi, N., B. Ruffner, et al. (2009). "High-level molecular diversity of copper-zinc superoxide dismutase genes among and within species of arbuscular Mycorrhizal fungi." Appl Environ Microbiol **75**(7): 1970-1978.

Croll, D., M. Giovannetti, et al. (2009). "Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus Glomus intraradices." New Phytol **181**(4): 924-937.

Croll, D. and I. R. Sanders (2009). "Recombination in Glomus intraradices, a supposed ancient asexual arbuscular mycorrhizal fungus." BMC Evol Biol **9**: 13.

Du, Z., X. Zhou, et al. (2010). "agriGO: a GO analysis toolkit for the agricultural community." Nucleic Acids Research **38**(Web Server issue): W64-70.

Dyer, P. S. and M. Paoletti (2005). "Reproduction in Aspergillus fumigatus: sexuality in a supposedly asexual species?" Med Mycol **43 Suppl 1**: S7-14.

E Tisserant, A. K., P Dozolme-Seddas, R Balestrini, K Benabdellah, A Colard, D Croll, C Da Silva, SK Gomez, R Koul, N Ferrol, V Fiorilli, D Formey, Ph Franken, N Helber, M Hijri, L Lanfranco, E Lindquist, Y Liu, M Malbreil, E Morin, J Poulain, H Shapiro, D van Tuinen, A Waschke, C Azcón-Aguilar, G Bécard, Bonfante P, MJ Harrison, P Lammers, U Paszkowski, N Requena, SA Rensing, C Roux, IR Sanders, Y Shachar-Hill, G Tuskan, JPW Young, V Gianinazzi-Pearson and F Martin (2011). "The transcriptome of the arbuscular mycorrhizal fungus Glomus intraradices reveals functional tradeoffs in an obligate symbiont." New Phytolgist *(Submitted)*.

Ebel, R. C., G. E. Welbaum, et al. (1996). "Arbuscular mycorrhizal symbiosis and nonhydraulic: Signaling of soil drying in Vigna unguiculata (L) Walp." Mycorrhiza **6**(2): 119-127.

Feddermann, N., R. R. D. Muni, et al. (2010). "The PAM1 gene of petunia, required for intracellular accommodation and morphogenesis of arbuscular mycorrhizal fungi, encodes a homologue of VAPYRIN." Plant Journal **64**(3): 470-481.

Fiorilli, V., M. Catoni, et al. (2009). "Global and cell-type gene expression profiles in tomato plants colonized by an arbuscular mycorrhizal fungus." New Phytologist **184**(4): 975-987.

Fischer, K. S., J. Barton, et al. (2000). "Genomics and agriculture. Collaborations in rice." Science **290**(5490): 279-280.

Francis, R. and D. J. Read (1984). "Direct Transfer of Carbon between Plants Connected by Vesicular Arbuscular Mycorrhizal Mycelium." Nature **307**(5946): 53-56.

Frey, B. and H. Schuepp (1992). "Transfer of Symbiotically Fixed Nitrogen from Berseem (Trifolium-Alexandrinum L) to Maize Via Vesicular Arbuscular Mycorrhizal Hyphae." New Phytologist **122**(3): 447-454.

Gandolfi, A., I. R. Sanders, et al. (2003). "Evidence of recombination in putative ancient asexuals." Molecular Biology and Evolution **20**(5): 754-761.

Gange, A. C., V. K. Brown, et al. (2003). "Multitrophic links between arbuscular mycorrhizal fungi and insect parasitoids." Ecology Letters **6**(12): 1051-1055.

Garcia-Garrido, J. M. and J. A. Ocampo (2002). "Regulation of the plant defence response in arbuscular mycorrhizal symbiosis." Journal of Experimental Botany **53**(373): 1377-1386.

Gautier, L., L. Cope, et al. (2004). "affy--analysis of Affymetrix GeneChip data at the probe level." Bioinformatics **20**(3): 307-315.

Genre, A., M. Chabaud, et al. (2008). "Prepenetration apparatus assembly precedes and predicts the colonization patterns of arbuscular mycorrhizal fungi within the root cortex of both Medicago truncatula and Daucus carota." Plant Cell **20**(5): 1407-1420.

Genre, A., M. Chabaud, et al. (2005). "Arbuscular mycorrhizal fungi elicit a novel intracellular apparatus in Medicago truncatula root epidermal cells before infection." Plant Cell **17**(12): 3489-3499.

Gentleman, R. C., V. J. Carey, et al. (2004). "Bioconductor: open software development for computational biology and bioinformatics." Genome Biol **5**(10): R80.

Giovannetti, M., D. Azzolini, et al. (1999). "Anastomosis formation and nuclear and protoplasmic exchange in arbuscular mycorrhizal fungi." Applied and Environmental Microbiology **65**(12): 5571-5575.

Goff, S. A., D. Ricke, et al. (2002). "A draft sequence of the rice genome (Oryza sativa L. ssp. japonica)." Science **296**(5565): 92-100.

Gow, N. A. (2005). "Fungal genomics: forensic evidence of sexual activity." Curr Biol **15**(13): R509-511.

Graham, J. H. and L. K. Abbott (2000). "Wheat responses to aggressive and non-aggressive arbuscular mycorrhizal fungi." Plant and Soil **220**(1-2): 207-218.

Grime, J. P., J. M. L. Mackey, et al. (1987). "Floristic Diversity in a Model System Using Experimental Microcosms." Nature **328**(6129): 420-422.

Grunwald, U., W. Guo, et al. (2009). "Overlapping expression patterns and differential transcript levels of phosphate transporter genes in arbuscular mycorrhizal, Pi-fertilised and phytohormone-treated Medicago truncatula roots." Planta **229**(5): 1023-1034.

Guether, M., R. Balestrini, et al. (2009). "Genome-wide reprogramming of regulatory networks, transport, cell wall and membrane biogenesis during arbuscular mycorrhizal symbiosis in Lotus japonicus." New Phytol **182**(1): 200-212.

Guimil, S., H. S. Chang, et al. (2005). "Comparative transcriptomics of rice reveals an ancient pattern of response to microbial colonization." Proc Natl Acad Sci U S A **102**(22): 8066-8070.

Gutjahr, C., M. Banba, et al. (2008). "Arbuscular mycorrhiza-specific signaling in rice transcends the common symbiosis signaling pathway." Plant Cell **20**(11): 2989-3005.

Gutjahr, C., L. Casieri, et al. (2009). "Glomus intraradices induces changes in root system architecture of rice independently of common symbiosis signaling." New Phytol.

Harrison, M. J. (1997). "The arbuscular mycorrhizal symbiosis: An underground association." Trends in Plant Science **2**(2): 54-60.

Harrison, M. J. (1999). "Biotrophic interfaces and nutrient transport in plant fungal symbioses." Journal of Experimental Botany **50**: 1013-1022.

Harrison, M. J. (1999). "Molecular and cellular aspects of the arbuscular mycorrhizal symbiosis." Annual Review of Plant Physiology and Plant Molecular Biology **50**: 361-389.

Harrison, M. J., G. R. Dewbre, et al. (2002). "A phosphate transporter from Medicago truncatula involved in the acquisiton of phosphate released by arbuscular mycorrhizal fungi." Plant Cell **14**(10): 2413-2429.

Hellemans, J., G. Mortier, et al. (2007). "qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data." Genome Biol **8**(2): R19.

Hijri, M., D. Redecker, et al. (2002). "Identification and isolation of two ascomycete fungi from spores of the arbuscular mycorrhizal fungus Scutellospora castanea." Applied and Environmental Microbiology **68**(9): 4567-4573.

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**(2): 253-261.

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

Hodge, A., C. D. Campbell, et al. (2001). "An arbuscular mycorrhizal fungus accelerates decomposition and acquires nitrogen directly from organic material." Nature **413**(6853): 297-299.

Hohnjec, N., M. E. Vieweg, et al. (2005). "Overlaps in the transcriptional profiles of Medicago truncatula roots inoculated with two different Glomus fungi provide insights into the genetic program activated during arbuscular mycorrhiza." Plant Physiology **137**(4): 1283-1301.

Hosny, M., J. P. P. deBarros, et al. (1997). "Base composition of DNA from glomalean fungi: High amounts of methylated cytosine." Fungal Genetics and Biology **22**(2): 103-111.

Hosny, M., V. Gianinazzi-Pearson, et al. (1998). "Nuclear DNA content of 11 fungal species in Glomales." Genome **41**(3): 422-428.

Irizarry, R. A., B. M. Bolstad, et al. (2003). "Summaries of Affymetrix GeneChip probe level data." Nucleic Acids Research **31**(4): e15.

Izawa, T., N. Ishida, et al. (1996). "[Transposon tagging using transgenic plants]." Tanpakushitsu Kakusan Koso **41**(9): 1415-1421; discussion 1397-1418.

Jakobsen, I. and L. Rosendahl (1990). "Carbon Flow into Soil and External Hyphae from Roots of Mycorrhizal Cucumber Plants." New Phytologist **115**(1): 77-83.

Jakobsen, I. and L. Rosendahl (1990). "N_2 Fixation and Root Respiration in Pea - Effects of AM and P." Agriculture Ecosystems & Environment **29**(1-4): 205-209.

Johnson, N. C., J. H. Graham, et al. (1997). "Functioning of mycorrhizal associations along the mutualism-parasitism continuum." New Phytologist **135**(4): 575-586.

Judson, O. P. and B. B. Normark (1996). "Ancient asexual scandals." Trends in Ecology & Evolution **11**(2): A41-A46.

Kleckner, N. (1996). "Meiosis: how could it work?" Proc Natl Acad Sci U S A **93**(16): 8167-8174.

Klironomos, J. N., J. McCune, et al. (2000). "The influence of arbuscular mycorrhizae on the relationship between plant diversity and productivity." Ecology Letters **3**(2): 137-141.

Kloppholz, S., H. Kuhn, et al. (2011). "A Secreted Fungal Effector of Glomus intraradices Promotes Symbiotic Biotrophy." Curr Biol **21**(14): 1204-1209.

Koch, A. M., D. Croll, et al. (2006). "Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth." Ecology Letters **9**(2): 103-110.

Koch, A. M., G. Kuhn, et al. (2004). "High genetic variability and low local diversity in a population of arbuscular mycorrhizal fungi." Proceedings of the National Academy of Sciences of the United States of America **101**(8): 2369-2374.

Kuhn, G., M. Hijri, et al. (2001). "Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi." Nature **414**(6865): 745-748.

Lang, B. F. and M. Hijri (2009). "The complete Glomus intraradices mitochondrial genome sequence--a milestone in mycorrhizal research." New Phytol **183**(1): 3-6.

Liu, J., L. A. Blaylock, et al. (2003). "Transcript profiling coupled with spatial expression analyses reveals genes involved in distinct developmental stages of an arbuscular mycorrhizal symbiosis." Plant Cell **15**(9): 2106-2123.

Liu, J. Y., L. A. Blaylock, et al. (2004). "cDNA arrays as a tool to identify mycorrhiza-regulated genes: identification of mycorrhiza-induced genes that encode or generate signaling molecules implicated in the control of root growth." Canadian Journal of Botany-Revue Canadienne De Botanique **82**(8): 1177-1185.

Martin, F., V. Gianinazzi-Pearson, et al. (2008). "The long hard road to a completed Glomus intraradices genome." New Phytol **180**(4): 747-750.

Massoumou, M., D. van Tuinen, et al. (2007). "Medicago truncatula gene responses specific to arbuscular mycorrhiza interactions with different species and genera of Glomeromycota." Mycorrhiza **17**(3): 223-234.

Mcgonigle, T. P., M. H. Miller, et al. (1990). "A New Method Which Gives an Objective-Measure of Colonization of Roots by Vesicular Arbuscular Mycorrhizal Fungi." New Phytologist **115**(3): 495-501.

McGuire, I. C., J. E. Davis, et al. (2005). "Heterokaryon formation and parasexual recombination between vegetatively incompatible lineages in a population of the chestnut blight fungus, Cryphonectria parasitica." Mol Ecol **14**(12): 3657-3669.

Morton, J. B. and G. L. Benny (1990). "Revised Classification of Arbuscular Mycorrhizal Fungi (Zygomycetes) - a New Order, Glomales, 2 New Suborders, Glomineae and Gigasporineae, and 2 New Families, Acaulosporaceae and Gigasporaceae, with an Emendation of Glomaceae." Mycotaxon **37**: 471-491.

Munkvold, L., R. Kjoller, et al. (2004). "High functional diversity within species of arbuscular mycorrhizal fungi." New Phytologist **164**(2): 357-364.

Newman, E. I. and W. R. Eason (1993). "Rates of Phosphorus Transfer within and between Ryegrass (Lolium-Perenne) Plants." Functional Ecology **7**(2): 242-248.

O'Gorman, C. M., H. T. Fuller, et al. (2009). "Discovery of a sexual cycle in the opportunistic fungal pathogen Aspergillus fumigatus." Nature **457**(7228): 471-474.

Okazaki, K. and O. Niwa (2008). "Dikaryotic cell division of the fission yeast Schizosaccharomyces pombe." Biosci Biotechnol Biochem **72**(6): 1531-1538.

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

Paszkowski, U., S. Kroken, et al. (2002). "Rice phosphate transporters include an evolutionarily divergent gene specifically activated in arbuscular mycorrhizal symbiosis." Proc Natl Acad Sci U S A **99**(20): 13324-13329.

Pawlowska, T. E. and J. W. Taylor (2004). "Organization of genetic variation in individuals of arbuscular mycorrhizal fungi." Nature **427**(6976): 733-737.

Peng, S. B., D. M. Eissenstat, et al. (1993). "Growth Depression in Mycorrhizal Citrus at High-Phosphorus Supply - Analysis of Carbon Costs." Plant Physiology **101**(3): 1063-1071.

Reddy, D. M. R. S., M. Schorderet, et al. (2007). "A petunia mutant affected in intracellular accommodation and morphogenesis of arbuscular mycorrhizal fungi." Plant J **51**(5): 739-750.

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

Remy, W., T. N. Taylor, et al. (1994). "4-Hundred-Million-Year-Old Vesicular-Arbuscular Mycorrhizae." Proceedings of the National Academy of Sciences of the United States of America **91**(25): 11841-11843.

Requena, N., P. Mann, et al. (2002). "Early developmentally regulated genes in the arbuscular mycorrhizal fungus Glomus mosseae: identification of GmGIN1, a novel gene with homology to the C-terminus of metazoan hedgehog proteins." Plant and Soil **244**(1-2): 129-139.

Ricken, B. and W. Hofner (1996). "Effect of arbuscular mycorrhizal fungi (AMF) on heavy metal tolerance of alfalfa (Medicago sativa L) and oat (Avena sativa L) on a sewage-sludge treated soil." Zeitschrift Fur Pflanzenernahrung Und Bodenkunde **159**(2): 189-194.

Rillig, M. C. (2004). "Arbuscular mycorrhizae and terrestrial ecosystem processes." Ecology Letters **7**(8): 740-754.

Salzer, P., A. Bonanomi, et al. (2000). "Differential expression of eight chitinase genes in Medicago truncatula roots during mycorrhiza formation, nodulation, and pathogen infection." Mol Plant Microbe Interact **13**(7): 763-777.

Sanders, I. R., M. Alt, et al. (1995). "Identification of Ribosomal DNA Polymorphisms among and within Spores of the Glomales - Application to Studies on the Genetic Diversity of Arbuscular Mycorrhizal Fungal Communities." New Phytologist **130**(3): 419-427.

Sanders, I. R., J. P. Clapp, et al. (1996). "The genetic diversity of arbuscular mycorrhizal fungi in natural ecosystems - A key to understanding the ecology and functioning of the mycorrhizal symbiosis." New Phytologist **133**(1): 123-134.

Sasaki, T. and B. Burr (2000). "International Rice Genome Sequencing Project: the effort to completely sequence the rice genome." Curr Opin Plant Biol **3**(2): 138-141.

Schussler, A., D. Schwarzott, et al. (2001). "A new fungal phylum, the Glomeromycota: phylogeny and evolution." Mycological Research **105**(12): 1413-1421.

Selosse, M. A. and F. Le Tacon (1998). "The land flora: a phototroph-fungus partnership?" Trends in Ecology & Evolution **13**(1): 15-20.

Shachar-Hill, Y., P. E. Pfeffer, et al. (1995). "Partitioning of Intermediary Carbon Metabolism in Vesicular-Arbuscular Mycorrhizal Leek." Plant Physiol **108**(1): 7-15.

Smith, J. (1998). Evolutionary Genetics, {Oxford University Press, USA}.

Smith, S. E. and D. J. Read (2008). Mycorrhizal Symbiosis.

Smyth, G. (2005). limma: Linear Models for Microarray Data. Bioinformatics and Computational Biology Solutions Using R and Bioconductor. R. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry and S. Dudoit, Springer New York: 397-420.

Stockinger, H., M. Kruger, et al. (2010). "DNA barcoding of arbuscular mycorrhizal fungi." New Phytol **187**(2): 461-474.

Team, R. D. C. (2008). "R: A Language and Environment for Statistical Computing. Vienna Austria."

Temnykh, S., W. D. Park, et al. (2000). "Mapping and genome organization of microsatellite sequences in rice (<i>Oryza sativa </i> L.)." TAG Theoretical and Applied Genetics **100**(5): 697-712.

Tommerup, I. C. and K. Sivasithamparam (1990). "Zygospores and asexual spores of Gigaspora decipiens, an arbuscular mycorrhizal fungus." Mycological Research **94**(7): 897-900.

Valentine, A. J., B. A. Osborne, et al. (2001). "Interactions between phosphorus supply and total nutrient availability on mycorrhizal colonization, growth and photosynthesis of cucumber." Scientia Horticulturae **88**(3): 177-189.

van der Heijden, M. G. A., T. Boller, et al. (1998). "Different arbuscular mycorrhizal fungal species are potential determinants of plant community structure." Ecology **79**(6): 2082-2091.

van der Heijden, M. G. A., J. N. Klironomos, et al. (1998). "Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity." Nature **396**(6706): 69-72.

van der Heijden, M. G. A., A. Wiemken, et al. (2003). "Different arbuscular mycorrhizal fungi alter coexistence and resource distribution between co-occurring plants (vol 157, pg 569, 2003)." New Phytologist **158**(3): 601-601.

Viera, A. and M. G. Glenn (1990). "DNA Content of Vesicular-Arbuscular Mycorrhizal Fungal Spores." Mycologia **82**(2): 263-267.

Weidmann, S., L. Sanchez, et al. (2004). "Fungal elicitation of signal transduction-related plant genes precedes mycorrhiza establishment and requires the dmi3 gene in Medicago truncatula." Molecular Plant-Microbe Interactions **17**(12): 1385-1393.

Wulf, A., K. Manthey, et al. (2003). "Transcriptional changes in response to arbuscular mycorrhiza development in the model plant Medicago truncatula." Mol Plant Microbe Interact **16**(4): 306-314.

Yamamoto, T., H. Lin, et al. (2000). "Identification of heading date quantitative trait locus Hd6 and characterization of its epistatic interactions with Hd2 in rice using advanced backcross progeny." Genetics **154**(2): 885-891.

Yokota, K. and M. Hayashi (2011). "Function and evolution of nodulation genes in legumes." Cell Mol Life Sci **68**(8): 1341-1351.

Annexe 1

Gene Copy Number Polymorphisms in an Arbuscular Mycorrhizal Fungal Population

Nicolas Corradi, Daniel Croll, Alexandre Colard, Gerrit Kuhn, Martine Ehinger, and Ian R. Sanders*

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

This chapter was published 2007 in *Applied and Environmental Microbiology*

Abstract

Gene copy number polymorphism was studied in a population of the arbuscular mycorrhizal fungus *Glomus intraradices* by using a quantitative PCR approach on four different genomic regions. Variation in gene copy number was found for a pseudogene and for three ribosomal genes, providing conclusive evidence for a widespread occurrence of macromutational events in the population.

Introduction

For a long time, gene duplication events were considered to occur relatively rarely among closely related genomes and it was assumed that individuals of extant populations are unlikely to differ in gene copy number. This is surprising given that the presence of gene copy number and length polymorphism in ribosomal DNA (rDNA) genes and pathogen and herbivore resistance genes has long been recognized to occur both within and among populations of fruit flies (*Drosophila melanogaster* and *Drosophila mercatorum*) and wild barley (*Hordeum vulgare*) (Coen *et al.*, 1982; Williams *et al.*, 1985; Zhang *et al.*, 1990). Recently, large-scale among-individual copy number polymorphisms for coding genes have been discovered in humans by surveying completely sequenced genomes or by using molecular hybridization techniques (Iafrate *et al.*, 2004; Sebat *et al.*, 2004). In some cases, these genomic changes have been linked to the susceptibility of individuals to pathologies such as human immunodeficiency virus or glomerulonephritis (Gonzalez *et al.*, 2005; Aitman *et al.*, 2006). In experimental populations of *Candida albicans*, copy number polymorphisms potentially evolved in as little as 330 generations (Cowen *et al.*, 2000).

Arbuscular mycorrhizal fungi (AMF) are important obligate symbionts of plants which improve plant nutrition and promote plant diversity in terrestrial ecosystems (van der Heijden *et al.*, 1998). Isolates of the AMF *Glomus intraradices* from one population were shown by fingerprinting to be genetically different (Koch *et al.*, 2004), and the same isolates were also shown to differentially affect plant biomass (Koch *et al.*, 2006). More recently, these *G. intraradices* isolates were found to harbor variable numbers of copies of genes that are potentially important in adaptation to environmental stress (P-type IID ATPases) (Corradi & Sanders, 2006). Considering that variation in gene copy number has been shown to be important for the adaptation to certain environmental or pathological conditions in other species, determining the presence and amount of such macromutational events in natural AMF populations is certainly warranted. From an evolutionary point of view it is also interesting to determine whether copy number polymorphisms are frequent enough within natural populations to be an important source of genetic variability. The issue which has been addressed here is whether variations in gene copy number are more widespread in a population of AMF or just restricted to the P-type IID ATPases. To address this question, we determined and compared the relative copy number of three rDNA genes (18S, 5.8S, and 25S)

and a *BiP* gene (a chaperone gene) and pseudogenes among isolates of *G. intraradices* harvested from the same field.

Results and discussion

Isolates of *G. intraradices* DAOM181602, A4, B3, C2, and C3 (the last four having been harvested from the same field in Switzerland and named according to the work of Koch *et al.* (2004)) were grown with Ri T-DNA-transformed *Daucus carrota* roots. DNA was extracted from AMF using the DNeasy plant minikit (QIAGEN). The *BiP* gene cloning procedure was performed according to the work of Kuhn *et al.* (2001). PCR amplification of specific variants of the *BiP* pseudogene presented in Fig. 1C was performed using the primer BiP2.F (5'-AAGACAAGCCACAAAAGATGCTGG-3') in combination with BiPT1.R (5'-TGAATATCATTGGTATATCCGTATATCT-3') or BiPT2.R (5'-TGAATGTCATTGG TATATCTCCGG-3'). Southern blot analyses were performed for the *BiP* genes of the isolate DAOM181602 using 3 µg of genomic DNA digested with each of the endonucleases EcoRV and XbaI. The hybridization was carried out using a digoxigenin-labeled probe obtained using the primers BiP2.F (5'-AAGACAAGCCACAAAAGATGCTGG-3') and BiP2.R (5'-AGTAGGGATTACAGTGTTACGAGG-3').

Real-time quantitative PCR (qPCR) was performed on three *G. intraradices* isolates to compare relative copy numbers of the *BiP*, 18S, 5.8S, and 25S genes. For each gene we designed primers and probes that annealed to a region that was conserved among all known sequence variants and among the isolates that we studied (Table 1). The control assays for real-time qPCR were performed according to the work of Corradi and Sanders (2006) and showed that cycle threshold (C_T) values did not vary significantly between independent DNA extractions. The average difference in C_T values among DNA extractions was <0.1 when qPCR was performed. In addition, control real-time qPCR was performed on the *Rad15* and *Rad32* genes. These two genes were chosen as controls as they showed no variation in sequence both among and within isolates and because *Rad32* had already been shown to be present in a single copy in isolate DAOM181602 by a dot blot hybridization assay (Hijri & Sanders, 2004). Probes were labeled with 6-carboxyfluorescein at the 5' end and 6-carboxytetramethylrhodamine at the 3' end (except for *Rad15* and *Rad32*, where Black Hole Quencher 1 was used at the 3' end). The 6-carboxyfluorescein-real-time PCR amplification was performed and the relative copy number of the genes was calculated according to the work of Corradi and Sanders (2006).

Three different variants of the *BiP* gene were found in the isolates DAOM181602, A4, B3, and C3. One of these variants harbored a frameshift mutation and should, therefore, be considered a pseudogene, while the two other sequence variants carried putative functional genes. The pseudogene of isolates A4 and C3 was found to harbor an additional frameshift mutation that did not occur in the other isolates. Isolate C2 was found to harbor both of the pseudogenes that were found in the other isolates (Fig. 1A). Therefore, we isolated four different *BiP* gene and pseudogene sequence variants from isolate C2, instead of three as for the other isolates. Consistent with the number of sequence variants isolated, Southern blotting on genomic DNA from isolate DAOM181602 confirmed that these genes are present in three different regions of the genome (Fig. 1B) and most likely appeared by gene duplications. PCR cloning with specific primers confirmed that isolate C2 harbored two *BiP* pseudogene variants, instead of one (Fig. 1C).

Relative quantification of *BiP* gene copy number was carried out on three *G. intraradices* isolates and showed that isolate C2 harbors, on average, 39% more copies of the *BiP* genes than do other isolates (Fig. 2A). This value is consistent with isolate C2 harboring four *BiP* genes instead of three. No significant variation in gene copy number was found between isolates B3 and C3. The search for copy number polymorphisms using qPCR in the *G. intraradices* population was extended to three rDNA genes (18S, 5.8S, and 25S). Estimates of the relative copy numbers of the three rDNA genes in isolates B3, C2, and C3 all showed the same pattern, with isolate C2 having the highest number of copies, followed by isolate B3, and with isolate C3 having the lowest number of copies. The differences in C_T value showed that in the *G. intraradices* population, the number of rDNA genes can vary from two- to fourfold among isolates (Fig. 2B to D). If the relative copy numbers of the three rRNA genes among isolates had been different from one another, then this could have represented independent duplication events. However, because all the rDNA genes showed the same relative differences among isolates, it is likely that the whole rDNA tandem array (comprising the three rRNA genes) has been subject to deletion and duplication events in the genomes of the three isolates. Using the same replicate DNA extractions from isolates B3, C2, and C3, we found no significant variation in C_T values (<0.1 C_T) for the amplification of either of the two control genes *Rad15* (Fig. 2E) and *Rad32* (data not shown).

In the present study we provide additional evidence that gene copy number polymorphisms can occur within an AMF population and that these involve several different regions of the genome. Together with the study by Corradi and Sanders (2006), this evidence shows that copy number polymorphism in this population occurs in rDNA, protein-encoding genes, and also pseudogenes. Obviously, copy number polymorphisms in protein-encoding genes could directly affect the phenotype. However, pseudogene expression is also known to regulate expression of homologous coding genes (Hirotsune *et al.*, 2003), and even changes in rDNA copy number can affect gene expression and epigenetic gene silencing (Michel *et al.*, 2005). Therefore, none of these copy number polymorphisms in the AMF population should be assumed to be neutral a priori, and these macromutational events could have potential consequences for the ecology of these fungi (Michel *et al.*, 2005).

Conclusion

Finally, the presence of copy number polymorphisms in AMF populations has direct implications for the study of AMF community structure. A main focus of many groups is currently to develop qPCR-based methods for assessing AMF community structure in plant roots, using the relative abundance of species-specific rDNA markers. The large within-population variation in rDNA gene copy number renders the results of such approaches largely uninterpretable.

Acknowledgements

This work was supported by a grant from the Swiss National Science Foundation (no. 3100A0-105790/1), and Gerrit Kuhn was supported by a grant from the Roche Foundation, from which support is gratefully acknowledged.

We thank all members of our research group for help in cultivation of *G. intraradices* isolates.

Figures and Tables:

Figure 1: Evidence for the segregation of *BiP* pseudogene variants among isolates of *G. intraradices*. A. Partial nucleotide sequence alignments of two *BiP* pseudogene variants isolated from five isolates of *G. intraradices* (DAOM181602, A4, B3, C2, and C3). The two different sequence types are shown in separate boxes. The number following each isolate code represents the pseudogene sequence type. B. Southern blot hybridization on genomic DNA from the isolate DAOM181602 of a probe that hybridizes to all *BiP* gene and pseudogene variants. Results with genomic DNA digested with EcoRV and XbaI are shown. The sizes of the bands are shown on the right. C. PCR amplification of the two *BiP* pseudogene variants using specific primers with DNA from the different *G. intraradices* isolates. The analysis shows results from an additional isolate (D1) from the same population that was not included in the cloning and sequencing experiments. The upper gel shows amplification of the *BiP* pseudogene type 1, using the reverse primer BiPT1.R. The lower gel shows amplification of the *BiP* pseudogene type 2, using the reverse primer BiPT2.R. The isolate C2 is shown to harbor both types 1 and 2 of the *BiP* pseudogene. The size standard is in the far left and far right lanes of both gels.

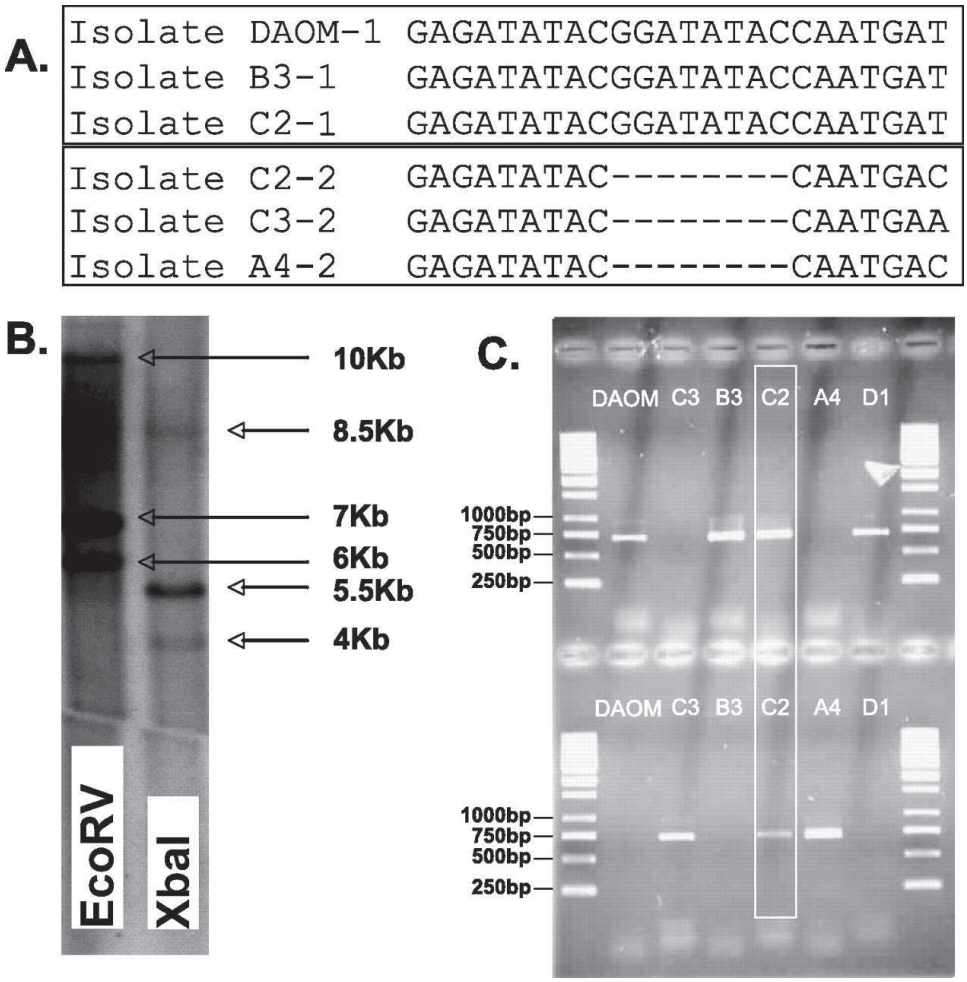


Figure 2: Results of real-time quantitative PCR showing linear regressions of the cycle threshold (C_T values) and the log concentration of genomic DNA of *G. intraradices* isolates B3, C2, and C3. The analysis was performed with primers amplifying all previously found variants of the *BiP* gene and pseudogenes (A); ribosomal genes 18S (B), 5.8S (C), and 25S (D); and *Rad15* (E). Real-time quantitative PCR was also performed on *Rad32*, but the data are not shown as they showed exactly the same pattern as that of *Rad15*. For all real-time quantitative PCR experiments, two replicate amplifications for each isolate were performed. Data points shown in the graphs represent the average C_T value of the two replicates.

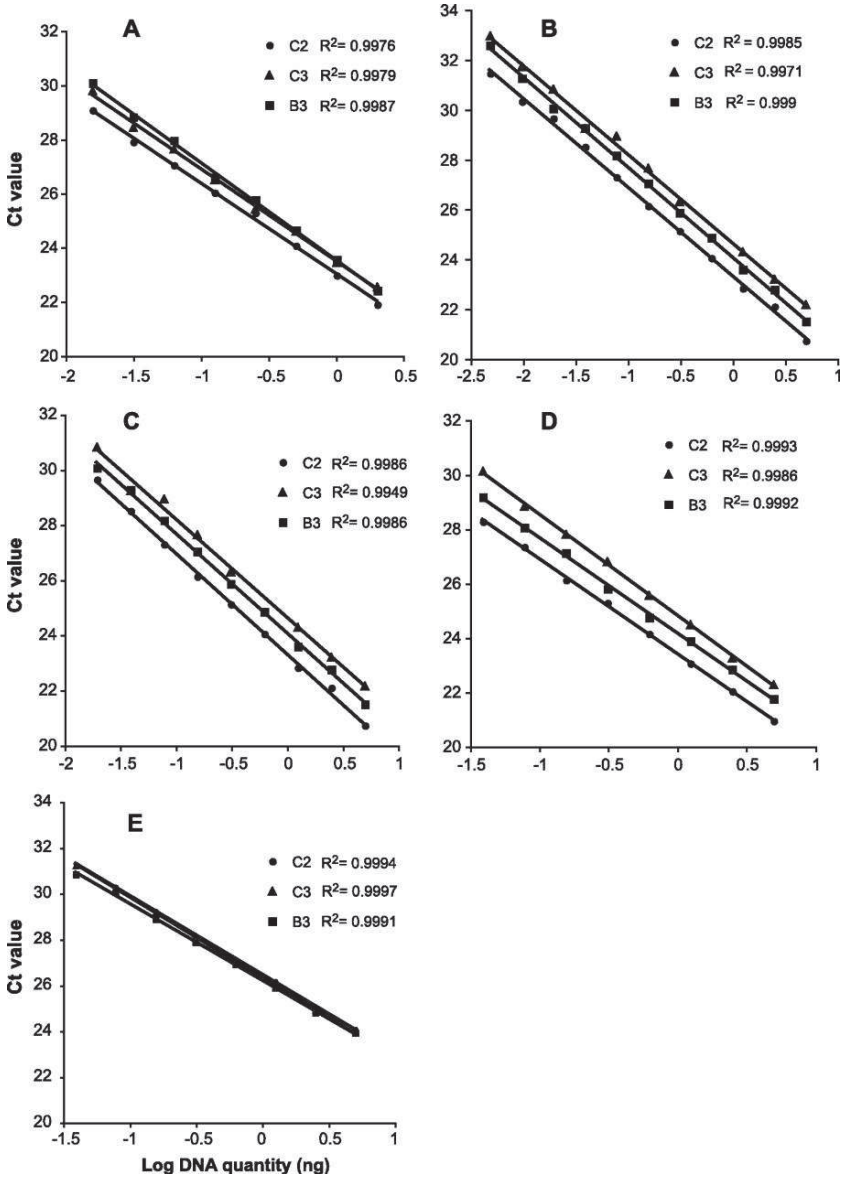


Table 1: List of primers and probes used in the real-time quantitative PCR experiments for amplification of BiP genes and pseudogenes; 18S, 5.8S, and 25S ribosomal genes; and the Rad15 and Rad32 genes

Primer names	Sense	5'-3' sequence
5.8S.real.F	Forward	ACA ACG GAT CTC TTG GCT CT
5.8S.probe	Probe	TGA AGA ACG TAG CGA AGT GCG
5.8S.real.R	Reverse	AAT TTG CGT TCA AAG ATT CG
18S.real.F	Forward	ATG CTA AAA CCT CCG ACT TC
18S.probe	Probe	TGA TTC ATA ATA ACT TTT CGA ATC GTA TGA
18S.real.R	Reverse	TAG GGC AGA AAT TTG AAT GA
25S.real.F	Forward	GAG AGA CCG ATA GCG AAC AA
25S.probe	Probe	CCG TGA GGG AAA GAT GAA AAG AA
25S.real.R	Reverse	GAA GGT ACG ACT GGC TTC AA
BiP.real.F	Forward	TTG GTG GTT CCA CAC GTA TTC
BiP.probe	Probe	ATG GTG CCG CCA TAC AAG
BiP.real.R	Reverse	TGT CAT AAC GCC ACC AGT TGT
Rad15 F	Forward	AGA GAG AAT TAT CGA ATA CGA GAA AAT GAT
Rad15 probe	Probe	TTT TGA CTT TTG ATG CCA TGC GCC AT
Rad15 R	Reverse	AGC CCA TAA TCT GTT TTC CCT CTT
Rad32 F	Forward	TCA ATA ATG ATC GCG ACT GAT ATA CA
Rad32 probe	Probe	TAT CAG GAG CAA GAT CCA ACA CGC GG
Rad32 R	Reverse	TAT TTC TCT GAA AGT GTT GAA GCT ATC G

Annexe II

High-Level Molecular Diversity of Copper-Zinc Superoxide Dismutase Genes among and within Species of Arbuscular Mycorrhizal Fungi

Nicolas Corradi,[1,2,‡] Beat Ruffner,[1,‡] Daniel Croll,[1] Alexandre Colard,[1] Ales Horák,[2], and Ian R. Sanders[1*]

Department of Ecology and Evolution, Biophore Building, University of Lausanne, CH-1015 Lausanne, Switzerland[1]

Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, British Columbia V6T 1Z4, Canada[2]

‡: N.C. and B.R. contributed equally to this work.

*: corresponding author

This chapter was published 2009 in *Applied and Environmental Microbiology*

Abstract

In the ecologically important arbuscular mycorrhizal fungi (AMF), *Sod1* encodes a functional polypeptide that confers increased tolerance to oxidative stress and that is upregulated inside the roots during early steps of the symbiosis with host plants. It is still unclear whether its expression is directed at scavenging reactive oxygen species (ROS) produced by the host, if it plays a role in the fungus-host dialogue, or if it is a consequence of oxidative stress from the surrounding environment. All these possibilities are equally likely, and molecular variation at the *Sod1* locus can possibly have adaptive implications for one or all of the three mentioned functions. In this paper, we analyzed the diversity of the *Sod1* gene in six AMF species, as well as 14 *Glomus intraradices* isolates from a single natural population. By sequencing this locus, we identified a large amount of nucleotide and amino acid molecular diversity both among AMF species and individuals, suggesting a rapid divergence of its codons. The *Sod1* gene was monomorphic within each isolate we analyzed, and quantitative PCR strongly suggest this locus is present as a single copy in *G. intraradices*. Maximum-likelihood analyses performed using a variety of models for codon evolution indicated that a number of amino acid sites most likely evolved under the regime of positive selection among AMF species. In addition, we found that some isolates of *G. intraradices* from a natural population harbor very divergent orthologous *Sod1* sequences, and our analysis suggested that diversifying selection, rather than recombination, was responsible for the persistence of this molecular diversity within the AMF population.

Introduction

Positive Darwinian selection is a potential signature of genetic conflicts. This has been demonstrated in host-pathogen interactions using molecular evidence (Andrews & Gojobori, 2004; Weber & Koebnik, 2006) and is also supported by theoretical predictions (Bishop *et al.*, 2000; Thrall *et al.*, 2001; Mondragon-Palomino *et al.*, 2002; Tellier & Brown, 2007). In sharp contrast, how genes evolve in a mutualistic environment remains largely unexplored. Strong selective constraints on genes that are specifically expressed during symbiotic stages are expected, and, indeed, slow rates of evolution have been predicted in conceptual models of mutualism because only weak selection pressures are expected in the host environment (Boucher *et al.*, 1982). Conversely, mutualism has been considered as a less harmonious interaction because the organisms involved have been predicted to evolve mechanisms to avoid overexploitation by their partners (Herre *et al.*, 1999; Kiers *et al.*, 2003). Evolution of such mechanisms in the host could create selective pressures for change in the symbiont. In the case of relatively unspecific mutualisms, a symbiont may have to maintain the ability to colonize different hosts that present different environments and have different abilities to avoid exploitation. Furthermore, competition among symbionts to colonize a host could provide selection pressures for rapid evolution of the genes involved in recognition and establishment of symbiosis. Recently, positive selection has been observed in the plant gene *NORK* that is involved in the establishment of the nitrogen-fixing mutualism with rhizobia and the arbuscular mycorrhizal symbiosis with fungi (de Mita *et al.*, 2006). It is highly pertinent to study evolutionary forces shaping the diversity of genes that are up- or downregulated in the symbiosis. This is an important and necessary step to predict their functions and their potential role between the mutualists.

The copper-zinc superoxide dismutase (Sod1) is a ubiquitous metalloprotein catalyzing the dismutation of superoxide into molecular oxygen and hydrogen peroxide. In the plant, expression of this molecule is usually one of the earliest detectable plant responses to pathogen infection and a key element of the plant defense system (Apostol *et al.*, 1989). In AMF, *Sod1* encodes a functional polypeptide that scavenges reactive oxygen species (ROS) from metabolic processes (Cox *et al.*, 2003). Potentially, it could be environmentally induced (i.e., heavy metal contamination in soils), but this is currently unknown. Furthermore, *Sod1* has been suggested to be an essential component in the plant/fungus dialogue necessary to

reach functional and structural compatibility between the partners (Garcia-Garrido & Ocampo, 2002; Lanfranco *et al.*, 2005), but this has not been experimentally demonstrated. However, the gene was shown to be strongly expressed in the root during early stages of the plant-AMF symbiosis (Lanfranco *et al.*, 2005). Furthermore, evidence of its ROS-deactivating role in roots is supported by observations that hydrogen peroxide accumulates in the same tissue and that blocking of *Sod1* in mycorrhizal roots markedly reduces hydrogen peroxide accumulation. Evidence therefore suggests that the *Sod1* gene plays an important role for the fungus during establishment of symbiosis with the host plant, but its exact function remains unclear.

AMF are putative ancient asexuals, belonging to the fungal phylum *Glomeromycota* (Scwarzott *et al.*, 2001; Corradi *et al.*, 2004a), that form obligate symbiotic interactions with about 80% of land plants. AMF are known as important determinants of plant biodiversity, ecosystem variability, and productivity (van der Heijden *et al.*, 1998), yet very little is known about the evolution of their genes and genomes. Understanding the evolutionary fate of AMF genes, as well as studying their molecular divergence within and among species, can provide useful insights about the evolution of these ecologically important symbionts (Dale & Moran, 2006). In this study, homologous sequences encoding a copper-zinc superoxide dismutase protein (Sod1) have been isolated from five AMF species and a collection of *Glomus intraradices* isolates that were all initiated from single spores from one agricultural field to identify the level of molecular diversity within an AMF population. Codon-based maximum-likelihood procedures and recombination detection methods have been used to identify potential selective forces acting on these genes throughout their evolutionary history.

Material and methods

Origin of isolates, AMF cultivation, and DNA extraction.

G. intraradices single-spore isolates were established from a population located in Switzerland. The 17 isolates examined originate from a small field site (90 m by 110 m) located at Hausweid, Tänikon. For details on spatial arrangement see Koch *et al.* (2004). All isolates of *G. intraradices* (Tänikon population and within-population comparisons), isolate DAOM 181602 (interspecific comparisons), *Glomus* sp. (strain MUCL 43196), *Glomus proliferum* (MUCL 41827), *Glomus geosporum*, and *Gigaspora rosea* were clonally subcultured every 15 weeks in two-compartment plates with root-inducing transfer DNA-transformed carrot roots (Bécard & Fortin, 1988). Freshly isolated hyphae and spores of each isolate and species were separately dried overnight at 48°C and ground into a fine powder using a Retsch MM300 mixer mill (Retsch, GmbH). The DNA was extracted using a modified version of the Cenis method for fungal DNA extraction with an additional step of a 1:1 dilution with a solution of 24:1 of chloroform-isoamyl alcohol before the final precipitation to remove remaining impurities (Cenis, 1992).

Gene amplification and cloning.

A total of four specific and slightly degenerate primers were designed based on the published *Sod1* sequence of *Gigaspora margarita* (Lnfranco *et al.*, 2005). Primer sequences were the following: 5'-AAAGCTATTGCAGTTCTTAG-3' and 5'-CCCAATAACACCACAGGCAA-3'; 5'-GCAGTTITTTASWCCTGATAA-3' and 5'-TAACACCACAGGCAACACGA-3'.

DNA amplifications yielded one single fragment of the expected size with each pair of primers tested. The PCR products were separated on agarose gel and purified using a MinElute Gel Extraction Kit (Qiagen). The fragments were cloned with a TOPO TA Cloning Kit (Invitrogen, Inc.) following the manufacturers instructions. DNA sequences were determined using a version 3.1 Terminator cycler sequencing kit and separated on an ABI Prism 3100 genetic analyzer (Applied Biosystems).

Sequence analyses, phylogenetic reconstruction, detection of positive selection, and recombination.

Sequences were analyzed and aligned using the Vector NTI package (Informax, Oxford, United Kingdom). Chromatograms were carefully examined by eye. A phylogenetic framework was reconstructed based on all *Sod1* genes and internal transcribed spacer (ITS) sequences using the neighbor-joining and minimum-evolution methods implemented in MEGA, version 3.1 (Kumar *et al.*, 1994). The maximum-likelihood tree was computed under the WAG+G model of evolution, chosen according to the Akaike information criterion (as computed with ProtTest, version 1.4 (Abascal *et al.*, 2005)), using Phyml, version 3.0 Guindon & Gascuel, 2003). Bootstrap support was calculated from 1,000 replicates. Bayesian posterior probabilities were calculated using MrBayes, version 3.1.2, (Huelsenbeck & Ronquist, 2001) with priors, chain number, and temperature set to default values; the aamodelpr parameter was fixed to particular protein models of evolution chosen according to the Akaike information criterion. Two parallel Markov chains were run for 2 x 106 generations; every 100th tree was sampled, and the first 5 x 105 generations were omitted from topology and probability reconstruction.

Codon-based likelihood ratio tests (LRTs), implemented in the PAML package, were used to detect potential signatures of selection in our data. Comparing the number of nonsynonymous to synonymous substitutions per site (*dN*/*dS*) is typically used to detect selection at the molecular level. Positive selection is inferred by a *dN*/*dS* ratio (ω) exceeding 1, while purifying selection is characterized by a $\omega$of <1, and neutrality is indicated by a $\omega$of 1. To test hypotheses of molecular evolution, we used models that allow different ωrates at different sites: the null model M1a (assuming two site classes: sites under purifying selection and neutrally evolving sites) and the alternative model M2a (adding a third site class: sites under positive selection) (Yang, 1997; Nielsen, 1998). Comparison of the log-likelihood values under the two models reveals whether the model that allows for positively selected sites fits the data set better than the null model. The LRT compares twice the difference of the log-likelihood to a χ^2 distribution with a degree of freedom equal to the difference of free parameters in the corresponding models. Naïve and empirical Bayes calculations were used to analyze in more detail the site class of positive selection to identify particular sites under positive selection. The codon-based likelihood ratio tests have been performed among the six AMF species using the phylogenetic relationships represented in Fig. S1A in the supplemental

material. The *G. intraradices* sequence used in the species data set corresponds to the isolate DAOM 181602. Rearranging the positions of *G. proliferum*, *G. geosporum*, and *Glomus* sp. within the AMF phylogeny did not affect the results reported in this study. The population data sets have been analyzed using the phylogenetic relationships shown in Fig. S1B in the supplemental material. The presence of recombination has been investigated using the program RDP2 (Martin *et al.*, 2005) and default parameters. Recombination events have been investigated using an alignment of nucleotide sequences of the *Sod1* gene from the totality of AMF species and isolates.

Real-time PCR procedures.

Real-time PCR was performed on three *G. intraradices* isolates (A2, A4, and C2) to compare relative copy numbers of the *Sod1* genes. Using the *Sod1* sequences we obtained from the 14 *G. intraradices* isolates, we designed primers that amplify a 100-bp fragment and that annealed to a region conserved among all variants. The primers and probe were as follows: forward primer (5' to 3'), CGA TTA CAG AGT TGG GAC CAC TTA; reverse primer (5' to 3'), CGT TGG CGA TTT GGG TAA TG; probe FAM (6-carboxyfluorescein), (5' to 3'): AGC AAC TTT ACC ATC CGG GCC AGC. The probe was labeled with FAM at the 5' end and with Black Hole quencher 1 at the 3' end. In the real-time PCR, the fluorescence of the probe was measured at each cycle at the annealing phase of the reaction. Real-time PCR amplification with the FAM-labeled probe was performed in 40 µl containing 1x qPCR Mastermix (Eurogentec), a 0.5 µM concentration of each primer, 0.25 µM probe, and 10 µl of DNA at different concentrations. The thermal cycling conditions comprised an initial step of 50°C for 2 min and 94°C for 10 min, followed by 45 cycles of 94°C for 15 s and 55°C for 1 min. Fluorescence data were collected using an ABI Prism 7000 Sequence Detection System (SDS; Applied Biosystems). The SDS software then generated each real-time PCR profile after multicomponent analysis by plotting the log of the change in fluorescence (delta Rn, where Rn is the normalized reporter signal) versus cycle numbers. The cycle threshold (C_T) was determined by the SDS software as the fractional cycle number. This indicates where the fluorescence crosses an arbitrary threshold intersecting the signal curves in their exponential phases. In each experiment, twofold serial dilutions of *G. intraradices* genomic DNA from the different isolates (ranging between 5,000 and 39.06 pg) were included to generate C_T values. Two independent experiments were performed, each with eight concentrations of the genomic

DNA. In half of the experiments, the DNA of the *G. intraradices* isolates was obtained from two independent extractions, thus allowing the possibility of checking whether a bias occurred in the C_Ts because of DNA extraction procedures. The C_T values shown in Fig. 3 represent averages between the two independent experiments. All replicates offered very similar results and produced standard curves with regression coefficients (R^2) of >0.99. The relative copy number of the *Sod1* gene in each of the *G. intraradices* isolates could then be calculated easily by comparing the C_T values of each of the isolates for a given amount of genomic DNA. The slopes resulting from the analyses of C_T values relative to the log of the amount of genomic DNA were consistent and reproducible among isolates, and the efficiency of the quantitative PCR from genomic DNA was found to be more than 95%.

Nucleotide sequence accession numbers.

AMF sequences obtained in this study have been deposited in GenBank under accession numbers EU232637 to EU232655.

Results and discussion

Genetic variability among AMF zinc-copper superoxide dismutases.

To better understand the molecular evolution of the *Sod1* genes in AMF, we sequenced 360 bp of the locus (almost the entire open reading frame) from five different AMF species (in two genera), as well as the genes from 14 different *G. intraradices* isolates harvested from the same field. All the sequences obtained showed a high degree of homology to the deposited sequences encoding fungal Cu-Zn superoxide dismutases according to a tBlastx search procedure (Altschul *et al.*, 1997). A phylogeny reconstructed using amino acid sequences from a variety of fungal Cu-Zn superoxide dismutase (*Sod1*) sequences, including the AMF sequences analyzed in this study, revealed the monophyly of all major fungal phyla, including the *Glomeromycota* (Fig. 1). The relationships we have identified between the fungal phyla are consistent with the most recent phylogenetic studies of the fungal kingdom (James *et al.*, 2006). Interestingly, all the phylogenetic models used resulted in a basal positioning of the chytrids and the zygomycetes within the fungal kingdom; however, their monophyletic origin was not supported by maximum-likelihood and Bayesian methodologies. Overall, statistical support for the major fungal clades was relatively low, a likely consequence of high amino acid sequence diversity within and across the major fungal phyla and the small size of the studied locus.

Amino acid comparisons among all the AMF species and isolates revealed a substantially low average sequence similarity of 73% among the analyzed species. This suggested that these sequences might have evolved either under neutrality or positive selective constraints for some time. High sequence variation was also observed among 14 *G. intraradices* isolates harvested from one field. A total of five different sequences of *Sod1* were obtained from 14 isolates of *G. intraradices*. Sequences from three of the isolates (A2, A4, and C2) accounted for most of the sequence variation, suggesting that the *Sod1* locus diverged significantly in these isolates, a feature that is unexpected for individuals from a single population (Table 1).

This surprisingly high molecular diversity warranted independent controls to ensure that it was not caused by an erroneous species assignment and to possibly rule out paralogous origin of the *Sod1* sequences from our population. These controls were performed as follows. All isolates of *G. intraradices* used in this study were identified through their ITS sequences, and

their monophyly was confirmed by using the closest related species, *Glomus diaphanum* and *Glomus fasciculatum* (Fig. 2 and Croll *et al.*, 2008b). In parallel, the putative orthology of the *Sod1* sequences and the absence of intraindividual variation were determined using two approaches. First, we screened and sequenced 32 clones of the *Sod1* gene in every single isolate used in this study, using two independent sets of primers (see Materials and Methods); no sequence variation was found within isolates. Based on the sequencing results, we were very confident that the molecular diversity we identified was representative of variation among isolates and that only a single sequence variant was present in each isolate. The absence of intraindividual variation also suggested that the divergent *Sod1* sequences did not arise through gene duplication (i.e., absence of paralogy). However, as cloning techniques are prone to randomly missing certain sequences and do not provide information about *Sod1* copy numbers, we performed a real-time PCR experiment using an additional set of primers (see Materials and Methods) targeting a highly conserved region of the *Sod1* gene in three of the most divergent isolates used in this study (Fig. 3). A significant difference in standard curves (one C_T difference corresponding to a twofold increase in gene copy number) would have suggested a difference in copy number for the *Sod1* gene among the isolates (Corradi & Sanders, 2006; Corradi *et al.*, 2007). The real-time PCR procedure showed very similar standard curves for all three isolates, suggesting that all *G. intraradices* isolates share the same copy number for the *Sod1* gene.

Potential occurrence of positive selection on AMF Sod1 genes.

Considering the amount of sequence variation we found among AMF species for the *Sod1* locus, we tested whether this locus might have evolved either under neutrality or under the regime of positive selection. To test this, we analyzed the *Sod1* sequences from the six different AMF species analyzed in this study using different models of codon evolution implemented in the PAML software package (Nielsen, 1998).

The results from the maximum-likelihood analysis of codon evolution showed that positive selection potentially acted at specific codons throughout the evolution of these AMF species. Indeed, all models that allowed for sites with a $\omega(dN/dS)$ greater than 1 (M1, M2, and M8), that is, the models of positive selection, fitted the *Sod1* data significantly better than the corresponding neutral models (M0, M3, and M7) (Table 2). More specifically, the ω estimates,

averaged over all sites and all branches, ranged from 0.5 to 0.61 among selection models. Models M1a and M7 gave lower estimates as they do not account for sites under positive selection. An LRT comparing model M1a (nearly neutral) to M2a (positive) indicated that the selection models fitted better (P = 0.009) (Table 3). The LRT comparing the model M7 (β, neutral) with M8 (β, selection) also rejected the null model (P = 0.004). Most sites in the fragment examined have apparently evolved either under purifying selection (M2a, 69%; M8, 92%) or neutrality (M2a, 26%; M8, 8%) (Table 2). The discrete model associated 10% of the sites under positive selection. Similarly, M2a and M8 indicated positive selection at a small subset of sites (4.7% and 7.7%) located within the fragment analyzed (Table 2). In more detailed analyses, a Bayesian approach identified seven sites affected by positive selection within the AMF lineage under model M8 and five under model M2a (Table 3). Overall, five amino acid sites were identified as potentially affected by positive selection when all sites that allowed for the presence of positive selection were used in the analysis (5G, 55F, 58T, 79P, and 88D). Sites 79P and 93D were predicted to be subject to positive selection at the 95% confidence level using all models.

In the Cu-Zn superoxide dismutase protein, the binding of copper and zinc ions requires the presence of histidine amino acids at specific conserved sites, and, as expected, these motifs did not show significant signs of positive selection. Apparently, the motifs potentially subject to positive selection are spread throughout the coding sequence, and none have been identified within the conserved internal catalytic cleft. Although the posterior probability for these mentioned sites to be under positive selective constraints was relatively high, we could not match any of these with a known function for the Sod1 protein. The significance of our tests might denote a functional importance for this rapid variation at the amino acid level, especially for the amino acids 55F and 58T, as these positions are located, respectively, 1 and 4 amino acids away from the to the catalytic cleft and have a 93% posterior probability (under model M3) of being shaped by events of positive selection.

Molecular diversity within a population of G. intraradices: recombination or diversifying selection?

Likelihood methods for detecting positive selection can be inaccurate if significant recombination has occurred among the sequences (Anisimova *et al.*, 2003). The *G.*

intraradices isolates analyzed in this study have been the subject of a multilocus genotyping study based on newly developed simple sequence repeat and nuclear gene intron markers (Croll *et al.*, 2008b). No events of recombination have been found between the isolates A2, A4, C2, and C5 using sequence data that was generated from these markers (Croll & Sanders, 2009). These isolates are the same ones that showed the highest molecular divergence in our study. To further test whether recombination played a significant role in shaping the diversity of the *Sod1* sequences identified in our population, we performed an analysis to detect possible recombination events within the *G. intraradices* population (Martin *et al.*, 2005). This analysis failed to identify any recombination events (data not shown), suggesting that other evolutionary forces played an important role in shaping the molecular diversity among the sequences.

To identify other possible evolutionary forces shaping the diversity at the *G. intraradices Sod1* locus, we applied the same likelihood procedures reported in the previous section to the 14 *G. intraradices* isolates (which were initiated as single-spore cultures) harvested from one field. Consistent with what we identified among AMF species, we found that selection models applied to our population explained the data significantly better than neutral models ($P < 0.05$) (Tables 2 and 3). Although an LRT comparison between the discrete model M1a with M3 was only marginally significant ($P = 0.09$), the more stringent comparisons using M2, M8, and M8A (supplementary category ω fixed at 1) models were significant ($P = 0.036$) (Table 3). Importantly, models accounting for positively selected sites produced very similar parameter estimates and associated the same proportion (4.9%) of amino acids to the supplementary category of positively selected sites. Testing alternative tree topologies to take into account the phylogenetic incertitude did not alter the outcome of our analysis. Overall, sites at positions 6 and 18 (6G and 18G) were predicted to be under positive selection by all models at a high level of significance. Importantly, amino acids 6G and 58F were commonly identified as being under positive selection at both the population and species levels (Table 3).

Selection on superoxide dismutases in fungal mutualists and pathogens.

AMF have sometimes been shown to harbor different alleles of one locus among the different genomes present in one isolate, but such intraindividual variability was not identified at the *Sod1* locus. Although it could be important for the AMF to have copies with various functions

for the *Sod1* gene, the absence of such allelic variation is not surprising and has also been reported for other adaptively important genes, such as the plasma membrane-type II ATPases (Corradi & Sanders, 2006).

Rapid evolution of *Sod1* has already been reported in pathogenic fungi (Chaturvedi *et al.*, 2001). For example, the Sod1 enzymes from different varieties of the human pathogen *Cryptococcus neoformans* and *Aspergillus* sp. show substantial biochemical and physical differences (Holdom *et al.*, 1996; Cox *et al.*, 2003). In *Aspergillus* species, amino acid variation in the N-terminal domain of Cu-Zn dismutases has been shown to increase thermostability. This possibly affects their virulence as increased thermostability and activity have been reported in the most virulent strains with the potential to develop aspergillosis (Holdom *et al.*, 1996). While the role of the *Sod1* genes in defense systems in plants, immune system in animals, and the pathogenesis of some fungi has been well characterized, their molecular variability has rarely been discussed in the context of adaptive evolution.

Our results are consistent with a rapid evolution of the *Sod1* gene, which has been shown to be specifically upregulated during mycorrhizal symbiosis and upon exposure to root exudates. However, its specific role during the establishment of the mycorrhizal symbiosis is still unclear. For instance, it is not known whether its expression is directly related to the plant-fungus dialogue or whether it is in response to a change in environmental conditions dictated by the plant defense system (i.e., expression required to counteract the ROS released by the host). Rapid evolution of genes in a plant-fungal interaction might also be an indication of arms races, but such arms races are usually seen in host-pathogen interactions (Bergelson *et al.*, 2001) and not expected to evolve in a mutualisitic symbiosis. Thus, an arms race scenario, although possible in the case of the *Sod1* gene, seems unlikely considering the generalist nature of the AMF. From complementation studies in yeast, it has been demonstrated that the AMF *Sod1* encodes an active polypeptide that effectively scavenges ROS (Lanfranco *et al.*, 2005). Moreover, this almost certainly occurs in AMF-colonized roots because suppression of *Sod1* activity reduces build-up of peroxidase in the roots (Lanfranco *et al.*, 2005). It seems evident, therefore, that one important role for AMF *Sod1* is the scavenging of ROS. This adaptive scenario implies that AMF are facing environmental situations similar to pathogens and, therefore, need to counteract the same plant defense barriers upon establishment of the mycorrhizal symbiosis. In this case, positive selection on the copper-zinc superoxide

dismutase protein is better explained as a direct consequence of adaptation directed toward scavenging ROS produced by the host.

In parallel, Sod1 proteins have also been suggested as playing an important role in the fungal-plant molecular dialogue (Garcia-Garrido & Ocampo, 2002; Lanfranco *et al.*, 2005) although conclusive evidence has not been provided for this. In this communication function of the Sod1 proteins, the rapid evolution of their amino acids could be related to the fact that tightly coevolved mutualisms are constantly exposed to nonmutualistic cheaters or parasites. As *Sod1* plays a key role for pathogens to avoid plant immune systems (production of ROS), the AMF might need to slightly change their biochemical properties accordingly to avoid their proteins' being recognized as of pathogen origin. An alternate explanation could be related to the poor specificity of AMF species toward their host. High rates of protein evolution facilitating host interactions could render AMF species more broad in their host range or allow them to rapidly colonize hosts in the presence of competitors.

Diversifying selection in an AMF population?

The molecular evidence we have gathered in this study points toward an orthologous origin of the *Sod1* sequences we identified within the *G. intraradices* population. Although it is true that we cannot conclusively reject a paralogous origin for a few of these, the absence of AMF isolates and species harboring more than one *Sod1* sequence and the impossibility to absolutely reject the monophyly of our *G. intraradices* sequences are both additional lines of evidence that make this latter evolutionary scenario far less parsimonious than the one suggesting an accumulation of mutations among orthologous sequences.

Intraspecific sequence polymorphism for a given gene has most often evolved as a consequence of random mutations accumulated by genetic drift. In this case, amino acid replacements are considered as neutral, with no effect or weak effects on protein function. An important departure from neutrality is observed when recombination produces molecular diversity or when positive selection fixes advantageous mutations. In the latter case, selective sweeps of favorable mutations lead to an excess of nonsynonymous over synonymous substitutions between species.

In this study, the molecular diversity identified among spores of *G. intraradices* is unlikely a consequence of recombination as other studies and our own analyses failed to identify recombination events. However, even if recombination did account for part of the intraspecific variation, this would not have significant consequences on our codon-based likelihood analyses for two independent reasons. First, rare events of recombination have been demonstrated as not interfering with the detection of positive selection (Anisimova *et al.*, 2003). Second, LRT comparisons using the more realistic model M7 against M8 were shown to be robust to false positives due to recombination. This is particularly the case if the model M7 allows the positive selection pressure to vary between 0 and 1 (not accounting for positive selection) and if the alternative model M8 allows an additional discrete class with a *dN*/*dS* that could be estimated to be >1 (and thus account for positive selection) (Anisimova *et al.*, 2004). In this study, the LRT comparison between these two models was the most significant one. This suggests that even under rare recombination in the population, diversifying selection is certainly a likely explanation for the diversity observed at the *Sod1* locus.

Within a population, directional selection is thought to shift allele frequency in favor of a newly derived mutation through a fitness advantage to certain individuals over others. Signals of positive selection are, therefore, rarely detectable at the population level. However, a changing environment, with variable selective pressures over time (i.e., interactions with multiple symbiotic partners), could lead to balanced polymorphism that is maintained within a single population. This feature is known for genes involved in plant defense systems, mating systems in fungi and plants, or immune response in animals, where a high level of polymorphism is favored and maintained through selection (Richman, 2000). Thus, rapid protein evolution and balancing selection at the locus *Sod1* may offer evolutionary advantages for some isolates compared to others in a heterogeneous environment, resulting in the maintenance of high protein divergence within the population.

Acknowledgments

We thank Lena Burri, John Wang, Kevin Carpenter, and four anonymous reviewers for their comments on this work. This work was supported by a Swiss National Science Foundation grant (3100AO-105790/1) to I.R.S.

Figures and Tables:

Figure 1: Phylogenetic analysis of the amino acid sequences of the fungal Cu-Zn superoxide dismutase gene *Sod1*. *Sod1* genes from five AMF species and 14 isolates of *G. intraradices* were obtained in this study and compared with previously published sequences from the *Ascomycota*, *Basidiomycota*, and *Zygomycota* and a sequence from *Gigaspora margarita* deposited in the GenBank. Values at nodes correspond to bootstrap support (1,000 bootstrap replicates) from neighbor-joining (first), minimum evolution (second), and maximum likelihood (third) and to posterior probabilities (fourth) from Bayesian analyses. Only bootstrap supports above 50 or posterior probabilities over 0.75 are shown. Scale bar represents 0.05 substitutions per site. The sequences from *Rhizopus oryzae*, *Batrachochytrium dendrobatids*, *Phycomyces blakesleeanus*, and *Allomyces macrogynus* were obtained from http://www.broad.mit.edu/ annotation/genome, http://genome.jgipsf.org/ Phybl1/Phybl1.home.html, and http://www.bch.umontreal. ca/pepdb/pepdb.html. Please note that *Sod1* sequences from isolate C2 and C5 differ only by synonymous substitutions and, therefore, share the same amino acid sequence.

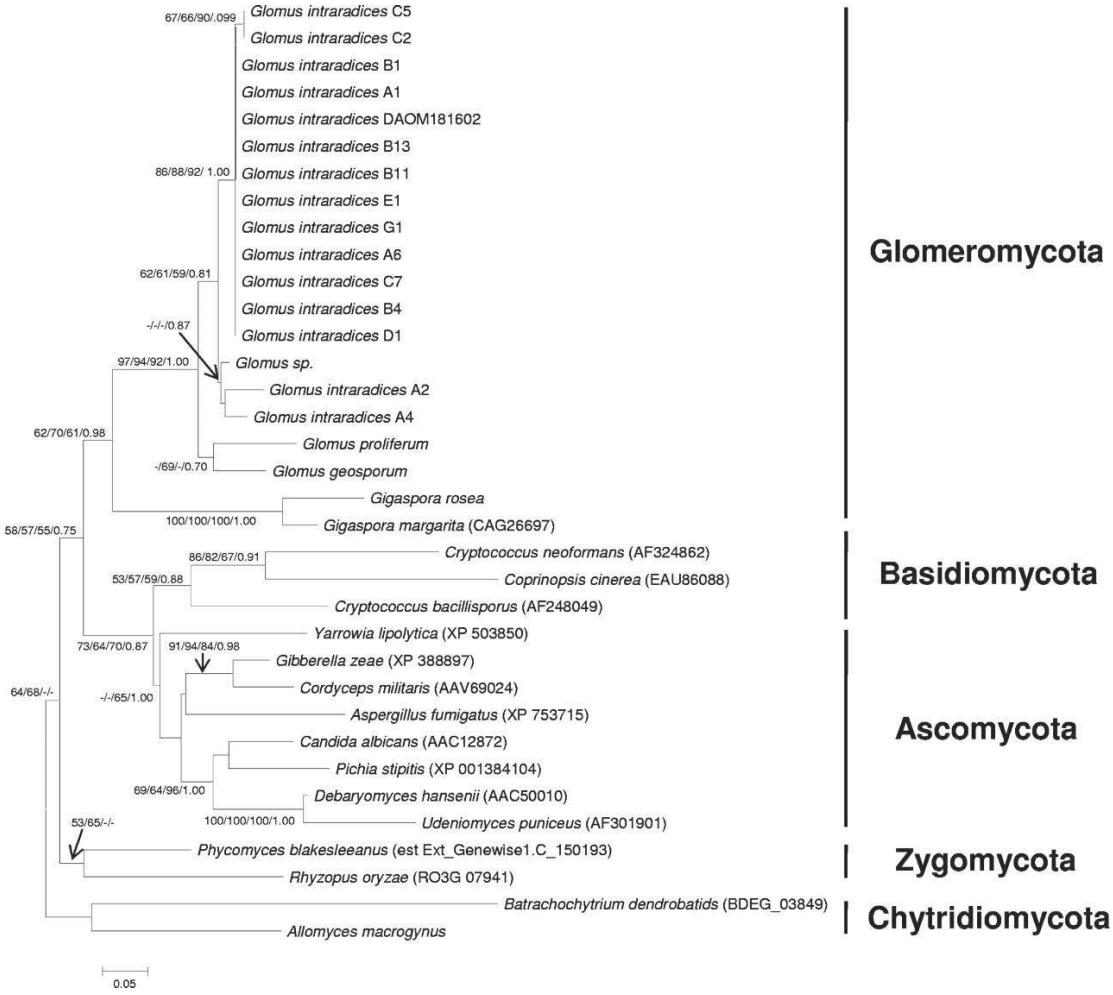


Figure 2: Consensus of neighbor-joining (uncorrected *P* distance) analyses is shown for ITS sequences from *G. intraradices* isolates analyzed in this study and harboring the most divergent *Sod1* sequences, as well as from *Glomus* sp. (strain MUCL 43196), *G. fasciculatum*, and *Glomus claroideum* (defined as the outgroup). Numbers on branches indicate clade support in the percentage of 1,000 bootstrap replicates (only clade support that is >75% is shown), using the neighbor-joining analysis. Sequences are identified by their GenBank accession numbers. Sequences in bold were recovered from isolates used in this study. Scale bar represents 0.02 substitutions per site. Our analyses suggest that the isolate MUCL 43196 likely corresponds to an isolate from *G. intraradices*.

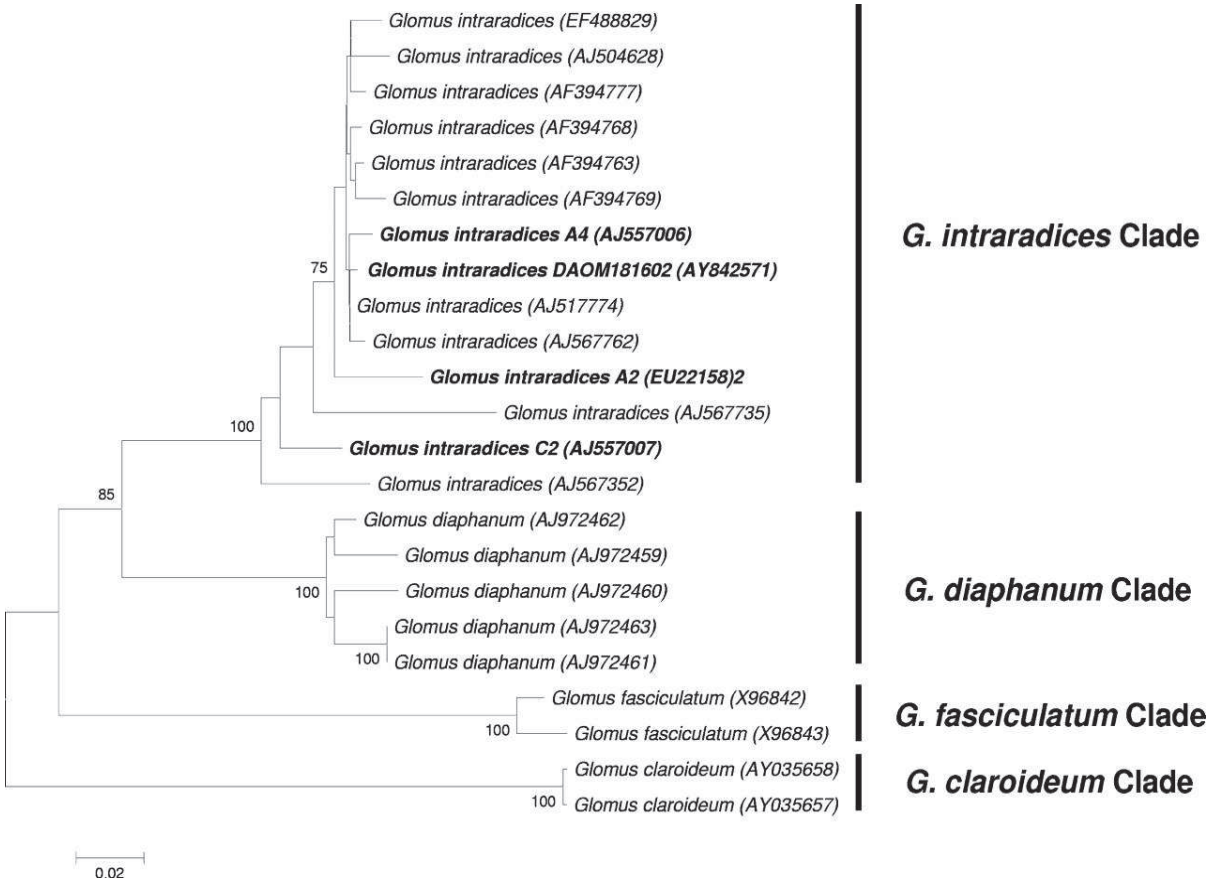


Figure 3: Results of real-time quantitative PCR showing linear regressions of the C_T values and the log concentration of the *G. intraradices* isolate A1, A2, A4, and C2 genomic DNA.

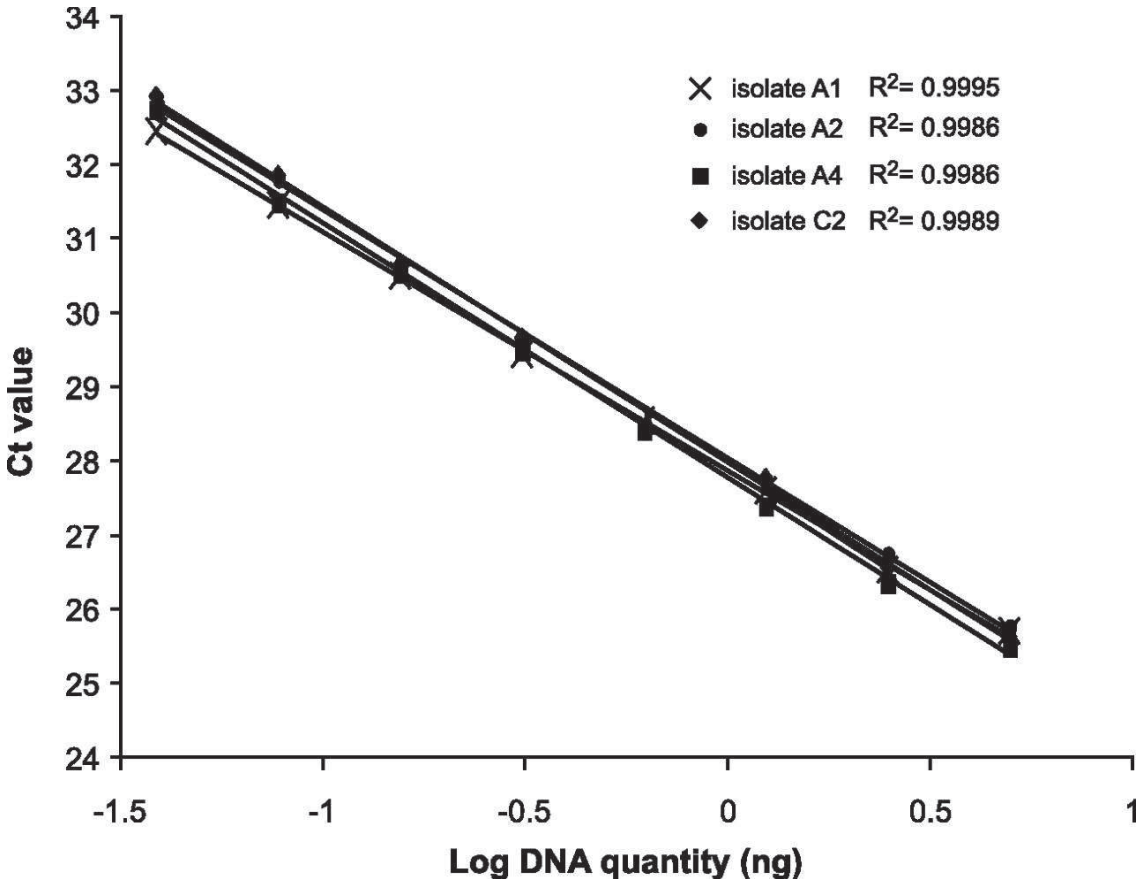


Table 1: Pairwise distances among representatives of the fungal Sod1 amino acid sequences analyzed in this study, distances were calculated using a Poisson correction (amino acid) included in the MEGA software, version 3

Glomus intraradices isolates																									
A1	0																								
A2	0.077	0.077																							
A4	0.0572	0.0572	0.067																						
A6	0	0	0.077	0.0572																					
B1	0	0	0.077	0.0572	0																				
B4	0	0	0.077	0.0572	0	0																			
B11	0	0	0.077	0.0572	0	0	0																		
B13	0	0	0.077	0.0572	0	0	0	0																	
C2	0.0093	0.0093	0.087	0.067	0.0093	0.0093	0.0093	0.0093	0.0093																
C5	0.0093	0.0093	0.087	0.067	0.0093	0.0093	0.0093	0.0093	0.0093	0															
C7	0	0	0.077	0.0572	0	0	0	0	0	0.0093	0.0093														
D1	0	0	0.077	0.0572	0	0	0	0	0	0.0093	0.0093	0													
E1	0	0	0.077	0.0572	0	0	0	0	0	0.0093	0.0093	0	0												
G1	0	0	0.077	0.0572	0	0	0	0	0	0.0093	0.0093	0	0	0											
Glomus sp. strain MUCL 43196	0.0187	0.0187	0.0572	0.0377	0.0187	0.0187	0.0187	0.0187	0.0187	0.0282	0.0282	0.0187	0.0187	0.0187	0.0187										
Glomus geosporum	0.1074	0.1074	0.1388	0.0972	0.1074	0.1074	0.1074	0.1074	0.1074	0.1178	0.1178	0.1074	0.1074	0.1074	0.1074	0.0972									
Glomus proliferum	0.1603	0.1603	0.1935	0.1603	0.1603	0.1603	0.1603	0.1603	0.1603	0.1713	0.1713	0.1603	0.1603	0.1603	0.1603	0.1713	0.1495								
Gigaspora rosea	0.4195	0.4195	0.4774	0.448	0.4195	0.4195	0.4195	0.4195	0.4195	0.4336	0.4336	0.4195	0.4195	0.4195	0.4195	0.448	0.448	0.4195							
Gigaspora margarita	0.3514	0.3514	0.4055	0.3781	0.3514	0.3514	0.3514	0.3514	0.3514	0.3646	0.3646	0.3514	0.3514	0.3514	0.3514	0.3781	0.3917	0.3781	0.1283						
Batrachochytrium dendrobatids	0.7118	0.7118	0.6931	0.7118	0.7118	0.7118	0.7118	0.7118	0.7118	0.7118	0.7118	0.7118	0.7118	0.7118	0.7118	0.6931	0.7503	0.8109	0.8535	0.8755					
Allomyces macrogynus	0.4626	0.4626	0.4774	0.4774	0.4626	0.4626	0.4626	0.4626	0.4626	0.4626	0.4626	0.4626	0.4626	0.4626	0.4626	0.4626	0.4626	0.5232	0.6568	0.5878	0.6568				
Phycomyces blakesleeanus	0.3254	0.3254	0.3781	0.3383	0.3254	0.3254	0.3254	0.3254	0.3254	0.3383	0.3383	0.3254	0.3254	0.3254	0.3254	0.3254	0.3646	0.3781	0.4925	0.448	0.7309	0.3917			
Rhyzopus oryzae	0.448	0.448	0.4774	0.4925	0.448	0.448	0.448	0.448	0.448	0.4336	0.4336	0.448	0.448	0.448	0.448	0.4774	0.5077	0.4774	0.6046	0.555	0.7309	0.5077	0.3383		
Coprinopsis cinerea	0.6748	0.6748	0.6391	0.6568	0.6748	0.6748	0.6748	0.6748	0.6748	0.6748	0.6748	0.6748	0.6748	0.6748	0.6748	0.6568	0.6568	0.6931	0.7503	0.6931	0.7503	0.6391	0.6391	0.6568	
Cryptococcus bacillisporus	0.448	0.448	0.4626	0.4195	0.448	0.448	0.448	0.448	0.448	0.4626	0.4626	0.448	0.448	0.448	0.448	0.4195	0.4336	0.4925	0.4774	0.4336	0.8109	0.555	0.4055	0.5077	0
Cryptococcus neoformans	0.5713	0.5713	0.6046	0.5713	0.5713	0.5713	0.5713	0.5713	0.5713	0.5878	0.5878	0.5713	0.5713	0.5713	0.5713	0.555	0.6046	0.6391	0.6391	0.6568	0.8755	0.7309	0.5232	0.6217	0
Candida albicans	0.3646	0.3646	0.4336	0.3917	0.3646	0.3646	0.3646	0.3646	0.3646	0.3781	0.3781	0.3646	0.3646	0.3646	0.3646	0.3781	0.4774	0.4626	0.5878	0.539	0.832	0.5713	0.4195	0.555	0
Aspergillus fumigatus	0.4626	0.4626	0.4774	0.4626	0.4626	0.4626	0.4626	0.4626	0.4626	0.4774	0.4774	0.4626	0.4626	0.4626	0.4626	0.4626	0.5077	0.5232	0.6391	0.5878	0.7903	0.5713	0.4195	0.5713	0
Debaryomyces hansenii	0.4055	0.4055	0.4055	0.4055	0.4055	0.4055	0.4055	0.4055	0.4055	0.4195	0.4195	0.4055	0.4055	0.4055	0.4055	0.3917	0.4626	0.5077	0.6046	0.5713	0.7503	0.6046	0.4774	0.539	0
Pichia stipitis	0.4336	0.4336	0.4925	0.4626	0.4336	0.4336	0.4336	0.4336	0.4336	0.448	0.448	0.4336	0.4336	0.4336	0.4336	0.4626	0.539	0.5232	0.5232	0.4925	0.8979	0.6046	0.4925	0.4925	0

Table 2: Site-specific models for positive selection, [a] Analysis was preformed using Codeml and site-specific models to test for positive selection in the AMF *Sod1* genes. [b] –ln*L*, log likelihood. [c] For ω, average estimates exceed 1 within the population of *G. intraradices* under the selection models, indicating positive selection. Among species, the estimated average ωranges between 0.5 and 0.6 but does not exceed 1. [d] p_0, p_1, and p_2 are the proportions of sites drawn from three discrete classes of ω. pand *q* are parameters of the β distribution.

Model and comparison (type)	–lnLb	ωc	Parameter(s)d
Within the G. intraradices population			
M0 (one ratio)	–562.07	0.898	None
M1a (nearly neutral)	–561.32	0.557	$\rho_0 = 0.443$ ($\rho_0 = 0.557$), $\omega_0 = 0.000$
M2a (positives election)	–558.00	1.938	$\rho_0 = 0.951$, $\rho_1 = 0.000$ ($\rho_2 = 0.049$), $\omega_0 = 0.476$, $\omega_2 = 30.241$
M3 (discrete)	–558.00	1.938	$\rho_0 = 0.000$, $\rho_1 = 0.951$ ($\rho^0 2 = 0.049$), $\omega_0 = 0.000$, $\omega_1 = 0.477$, $\omega_2 = 30.242$
M7 (β, neutral)	–561.35	0.5	$\rho = 0.005$, $q = 0.005$
M8 (β, positive selection)	–558.00	1.938	$\rho_0 = 0.951$ ($\rho_1 = 0.049$), $\rho = 90.348$, $q = 99.000$, $\omega = 30.249$
M8A (β, ω = 1)	–561.32	0.557	$\rho_0 = 0.443$ ($\rho_1 = 0.557$), $\rho = 0.005$, $q = 2.033$
Among species			
M0 (one ratio)	–1,165.38	0.201	None
M1a (nearly neutral)	–1,123.63	0.324	$\rho_0 = 0.697$ ($\rho_1 = 0.303$), $\omega_0 = 0.031$
M2a (positive selection)	–1,119.78	0.609	$\rho_0 = 0.680$, $\rho_1 = 0.275$ ($\rho_2 = 0.045$), $\omega_0 = 0.032$, $\omega_2 = 6.977$
M3 (Discrete)	–1,117.85	0.475	$\rho_0 = 0.538$, $\rho_1 = 0.371$ ($\rho_2 = 0.091$), $\omega_0 = 0.008$, $\omega_1 = 0.367$, $\omega_2 = 3.670$
M7 (β, neutral)	–1,123.49	0.271	$\rho = 0.144$, $q = 0.387$
M8 (β, positive selection)	–1,118.00	0.495	$\rho_0 = 0.928$ ($\rho_1 = 0.072$), $\rho = 0.197$, $q = 0.786$, $\omega = 4.291$
M8A (β, ω = 1)	–1,122.20	0.281	$\rho_0 = 0.766$ ($\rho_1 = 0.234$), $\rho = 0.308$, $q = 4.359$

Table 3: LRTs for positive selection, [a] LRTs of different models of ω variation among codons. [b] Selection models were preferred in all comparisons at the species level. In within-population comparisons, selection models M2a and M8 were preferential.[c] Twice the difference of the log likelihoods between two models. [d] df, degrees of freedom. [e] Sites possibly affected by events of positive selection, identified with a posterior probability of over 50%. Sites at a high level of significance are shown in bold. *, >95%; **, >99%.

Models comparedb	2ΔlnLc	dfd	*P* value	Positively selected sitese
Among AMF				
M0 vs M3	95.06	4	0.001	6 G*, 15 S, 55 F*, 58 T *, 63 K, 79 P**, 88 D**, 91 K
M1a vs M2a	11.56	3	0.009	6G, 55F, 58T, 79P**, 88D**
M7 vs M8	10.98	2	0.004	6G, 15S, 55F, 58T, 63K, 79P*, 88D*
M8 vs. M8A	8.4	1	0.004	
Within the G. intraradices population				
M0a vs M3	8.14	4	<0.09	6G*,18G*,58F
M1a vs M2a	6.64	2	0.036	6G*,18G*,58F
M7 vs M8	6.7	2	0.035	3 I, 6 G*, 14 E, 18 G*, 30 A, 38 I, 46 N, 49 T, 58 F, 61 T, 108 I
M8 vs M8A	6.64	1	0.01	

Annexe III

Conserved Meiosis-specific genes are present and expressed in the supposed asexual lineage Glomeromycota

Alexandre Colard, Nicolas Corradi and Ian R. Sanders

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

Abstract

The mutualistic arbuscular mycorrhizal fungi (AMF) are the biggest fungal phylum thought to be asexual. The genes involved in meiosis, mitosis and recombination are well known in other fungal organisms but no sequence information has been provided so far for AMF. By surveying EST and gene databases of the AMF *Glomus intraradices* we identified short sequences of several genes that could play a role in these processes. We increased the sampling in *Glomus intraradices* by designing several sets of degenerate primers for new genes playing role in meiosis, mitosis and recombination.

Consequently, we successfully isolated several genes of putative high importance in meiosis, mitosis and recombination from *Glomus intraradices*. By RT-PCR, we proved that these genes are expressed. We also show that these genes are under purifying selection.

This approach allowed us to identify the presence of gene families and events of gene duplications, as well as to identify the nature of selective forces that act on these genes in this putatively ancient asexual organism.

Introduction

Evolutionary theory predicts that ancient asexuals should not exist (Maynard, 1978). By definition, "sex" in eukaryotes is accomplished by meiosis, which is one of the most conserved processes in all kingdoms. Some genes are specific to the meiosis mechanism and are not found in other processes. A study describes a regrouping 7 meiotic core genes (Table 1 and Ramesh *et al.*, 2005). These genes are potentially sufficient for the meiosis process. Four of them are meiosis specific and are not expressed in other processes. The two other genes are also expressed in DNA repair and mitotic recombination. Even if the presence of meiosis specific genes or core meiotic genes is not evidence of meiosis itself, it would be surprising if asexual organisms would harbor these remarkably conserved genes. Giardia (Ramesh *et al.*, 2005), Entamoeba (Stanley *et al.*, 2005) and the microspordian Encephalitozoon (Gibson *et al.*, 2001) all harbor meiosis specific genes, even though there is no known sexual cycle, suggesting that undiscovered or cryptic sexual process may be present in these species.

Studies have already shown that recombination could have occurred in the AMF, *G. intraradices*. This was shown for the highly polymorphic BiP and rDNA sequences (Gandolfi *et al.*, 2003). However, because of the heterokariotic state of AMF, doubts still exist about whether this represents among-individual or among-nuclei recombination. A recent study using multi-locus sequence data of a *G. intraradices* population shows that recombination has occurred in a Swiss population of this fungus (Croll *et al.*, 2009). Nevertheless, meiosis has never been observed in an AMF. The aim of my study is to look for molecular evidence of meiotic recombination in the AMF *G.intraradices*. I used three different approaches. First, I used highly degenerate primers to try to amplify meiotic core genes. Second, I used sequences from the *Glomus* consortium and Joint Genome Institute to find genes. Third, I tested the expression of each meiosis specific gene.

Here, I provide evidence for the conservation of at least two meiosis-specific genes (*dmc1*, *msh4*,) in the genome of *G.intraradices*. Other genes that play role in meiosis, but not exclusively, were found, such as Mre11, Rad50 Rad51, pms1 and mlh1. I will discuss the possibilities which could explain this.

Genes	Function	Meiotic Core drive	Genes found in *Glomus intraradices*
Spo11*	Create double strand break (DBS) in DNA	*	
Dmc1*	Bind single DNA strain during crossing over	*	*
Msh4*	MutS homolog, required for reciprocal recombination and proper segregation of homologous chromosomes at meiosis	*	*
Msh5*	MutS homolog, facilitate crossovers between homologs during meiosis	*	*
Rad51	May participate in a common DNA damage response pathway associated with the activation of homologous recombination and double-strand break repair.	*	*
Rad50	Plays a central role in DSB, repair, DNA recombination, maintenance of telomere integrity and meiosis, RAD50 may be required to bind DNA ends and hold them in close proximity.	*	*
Rad32 (MRE11)	Component of the MRN complex, which plays a central role in double-strand break (DSB) repair, DNA recombination, maintenance of telomere integrity and meiosis. single-strand endonuclease activity and double-strand-specific 3'-5' exonuclease activity		*
Pms1	MutL homolog, postmeiotic segregation 1, involved in the repair of mismatches in DNA		*
Mlh1	MutL homolog, involved in the repair of mismatches in DNA	*	*

Table1: list of genes considered to play an important role in meiosis (Ramesh *et al.*, 2005). (*meiosis specific)

Material and methods

G. Intraradices cultivation and genomic DNA extraction

A culture of G. Intraradices (isolate DAOM 181602) growing on RiT-DNA transformed *Daucus carrota* roots was obtained from the group of Prof. G. Bécard (University Paul Sabatier, Toulouse, France). Fungi were grown without contamination on two compartment plates (St-Arnaud *et al.*, 1996). The cultures were grown for an average period of 4 months at 25°C before DNA extraction. The harvest was performed by dissolving medium containing AMF (Nagahashi *et al.*, 1999). Fungal material was then collected on a 22µm sieve.

The DNA was extracted using a modified version of the Cenis method for fungal DNA extraction (Cenis, 1992) with an additional step of 1:1 dilution with a solution of 24:1 of chloroform isoamyl alcohol before the final precipitation, to remove remaining impurities.

Sequence analysis

Blastx/ Blastp / tBlastx / tBlastn / PHIBlast were performed on NCBI web site. ClustalW (Thompson *et al.*, 1994) was used for sequence alignment and correction was doing with hands on Bioedit software (The BioEdit Sequence Alignment Editor software, Department of Microbiology, North California State University). Highly degenerate primers were designed from these alignments (Table 2).

Glomus sequences are regrouped on two web sites: *Glomus* Consortium web site (http://mycor.nancy.inra.fr/IMGC/GlomusGenome/index.html) and JGI website http://darwin.nmsu.edu/~plammers/). These websites were used to seek *Glomus* meiosis genes that we did not find by PCR amplification.

PCR Amplification, cloning and sequencing

Amplifications were carried out in a final volume of 25µl containing 1xPCR buffer (10mM Tris-HCL, 50 mM KCl, 1,5mM MgCl2, 0,1% Triton X-100 and 0,2mg/ml BSA), 2,5mM dNTP, 25µM final concentration of each degenerate primer, and 0,5U of Taq polymerase (Qbiogene©). PCR was performed in an automated thermal cycler (T-Gradient, Biometra). Reaction conditions were 94°C for 3min; 40 cycles of denaturation at 94°C for 30 s annealing at 42°C for 30 s and extension at 72°C for 1 m 30 s; final extension at 72°C for 7 m. Primers used are given in table 4. The Invitrogen© TOPO TA kit was used for cloning,

following the manufacturers instructions. Sequences were obtained using M13 forward and reverse primers with the Big Dye 3.1 Terminator cycle sequencing kit (AppliedBiosystems), according to the manufacturers' instructions, and separated on an ABI Prism 3100 genetic analyser (Applied Biosystems).

Sequence extension

The DNA Walking SpeedUp Kit with SeeAmp ACP PCR Master Mix II (from BioCat GmbH) was used for extending the sequences. Specific primers were designed using the Primer3 web site (http://frodo.wi.mit.edu/) and are given in Table 3

RNA Isolation

From same culture described above, we used the kit ToTally RNA Isolation from Ambion© and Promega© SV total RNA isolation following the instructions of the manufacturer.

RT-PCR

RT-PCR was performed with invitrogen© Superscript III. Primers are given in the table 4.

Phylogenetic analysis

Neighbor-joining and minimum evolution (Kumar *et al.*, 1996) genetic distances were calculated using MEGA4 (Kumar *et al.*, 2007) for both nucleotide and amino acid sequence data. Phylogenetic tree topologies were assessed by 1000 bootstrap replicates

Gene	Primer name	Primer sequences (5'->3')
Mlh1	F1	GAR AAY AGY XTX GAY GCN GG
	F2	AAY CAY TWY ACX AGY AAR T
	F3	GGN TTY CGX GGN GAR GC
	R3	GCY TCN CCX CGR AAN CC
	R4	AYR TTN ACR TCA
Mnd1	F1	TNA ARG ANY TIG ARA ARY T
	F2	ACN AGY AAY TAY TAY TGG
	F3	GAN CGI AAR GTR GGN AC
	R1	ATR TCN GTC CAN CKR TT
	R2	YIA YIT CYT TNA C
	R3	ATR TCN GTC CAN C
Msh4	F	AAY ATG AGY GGN AAR AGY
	R	CCY CTI CCN ARI TCR T
Msh5 F	F	TCI RTI TAY YTN AAR CAR GT
	R	CCY TTI CCR AAY TCR TCD AT
Red1	F1	CCI AAR ATH TCN GAR GTN CAR
	R1	DAT DAT YTT RTT CAT YTC
	F1	ATH GAR ACI GGN TCN
Spo11	F	CGN GAIN TNT AIT AIA
	F2	AAR MGI GAY ATD TWY TA
	405F3	AAI GGI YTI RTN GCN GG
	490F4	TIR TIR TNG ARA ARG C
	R	GCR TCY TTY TCY A
	R2	RTC NGG RTA NCC YTT NGG NCA
	620R3	IAR NCC NAR GGT
	575R4	IAD ICC RTR NGG RTC

Table 2: Highly degenerate primers used for PCR in order to find meiosis specific genes. F is for Forward primer; R is for Reverse primer. For each primers pair, a temperature gradient was performed between 40°C and 58°C to find optimal annealing conditions.

Gene	Primer name	Primer sequences (5'->3')
Mre11	TPS1-F	GGC AAA ATC AAT AAG GTG GA
	TPS2-F	TTC TGG GGA CAC GAA CAT GA
	TPS3-F	GCC AGG TT CTC CAG TGG CTA
Msh4	TPS1-R	ATG AGC AAA TGA CGTT AGC A
	TPS2-R	CGA TTG ATT CGG AAG CTT TG
	TPS3-R	GTT GTG AAG GTC AAA GTT TTC C
Rad50	TPS1-R	CGT CAA GAA CAA ACA TTC CA
	TPS2-R	CAT TCT GCT AAC GCC ATT CG
	TPS3-R	AAC ATC CTT TGG CCA GCA CT

Table 3: Primer design for the chromosome walking. Annealing temperature was given by the manufacturer

Gene	Primer name	Primer sequences (5'->3')
Dmc1	F	GAT TAT GCT GGT CGT GGT GA
	R	GCA AGT ACA TGA CCG CCA AT
	RT-F1	TCG CT G AAC GAC AAC AAA AG
	RT-R1	TTG CGT GGG TCA CTT ACA AA
	RT-F2	CTC GCT GAA CGA CAA CAA AA
	RT-R2	TTG CGT GGG TCA ACT TAC AAA
Mlh1	F1	GGT TAT AAT TCG CGT TGG AGA
	R1	CCA AAA GCA TGT GCT GGT AA
Msh4	F1	TCA TTT ACC GGG GTT CAA AA
	R1	TCA TCG TTT CTT CCG TCT TTG
	F2	CGC GAT TTA TTC GGT TTT GT
	R2	GCC TTT CTA ATA ACG TTT TCT GGT
	RT-F1	GGC TGG AAG AAA TCC AAT GA
	RT-R1	TCC GCT CAT ATT TGG ACC AG
	RT-F2	TTG GCA ATA AAG GCT GGA AG
	RT-R2	TCC GCT CAT ATT TGG ACC AG
Msh5	F2	AAG GCT GCA CAA AAT GTT CC
	R2	ATC AAT CGC GCT TCG TAA CT
Rad50	F	TGG TAT GCG TGG GAG ATG TA
	R	CGC AGG AAT TAC AAA GTT ATA TTC AA

Table 4: Primer design for the gene amplification on genomic DNA or on cDNA. (F: Forward primer; R: Reverse primer; RT : Primers designed specifically for the cDNA amplification) All these primers have an annealing temperature of 58°C.

Results

We found some sequences with a good blast score for two meiosis specific genes dmc1 and msh4 (2499pb for dmc1, 1230pb for msh4). We then used JGI data bases of the *G. intraradices* genome for scanning several genes playing a role in meiosis that had not been found using degenerate primer approach. Because of the degeneration of the genetic code, we used protein sequences of different fungi to perform tblastn and tblastx searches and found in those data some sequences matching Rad50, Rad51, mlh1, pms1 and meiosis-specific genes msh5. Nevertheless, the JGI database is not yet complete, and so it is possible that other genes are present in the *G.intraradices* genome. Moreover, these sequences are only partial, and no start or stop codons have been found. We focused more closely on dmc1 and msh4 because of their importance in meiosis and also because their sequences are longer than the others available. By alignment with genes of other species, we could define intron/exon topology for the two genes. There are ten exons for *Dmc1* and seven for *msh4*. Comparison of the protein sequences with S. *cerevisae* in the activity domains is about 77% and 65% for *dmc1* and *msh4* respectively. The GC content of the sequence is 30.09% for dmc1 and 30.45% for msh4; a similar value to those already found for the others genes in AMF (Corradi *et al.*, 2006).

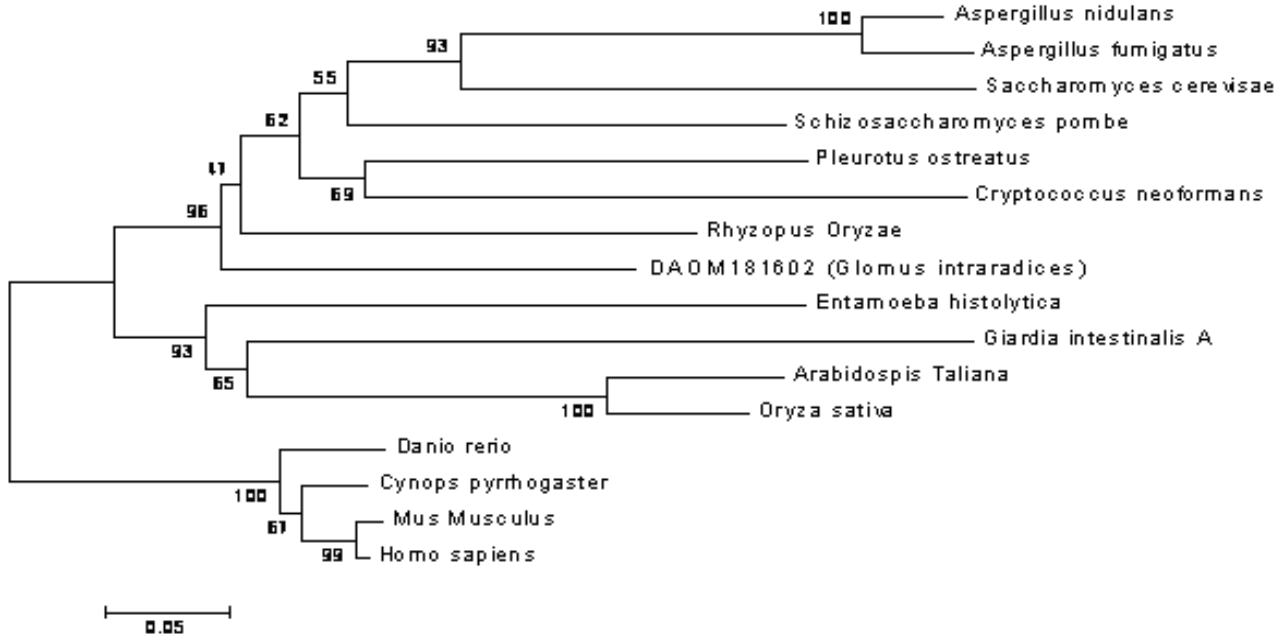


Figure 1: Phylogenetic analysis of meiosis specific Dmc1 (833 amino acid length). The Glomeromycota Dmc1 sequence was compared with previous published sequences from fungi, mammals, fishes, protist and plants. Numbers at nodes correspond to bootstrap support for neighbor-joining analyses. Scale bar represents 0.05 substitutions per site.

Phylogenetic analysis of the dmc1 and msh4 sequences using neighbor joining yielded trees in which we could see a clear distinction between each clade of fungi, ascomycota, basidiomycota, zygomycota and glomeromycota (Figure 1 and 2). This confirms that sequences are unlikely to be contaminations

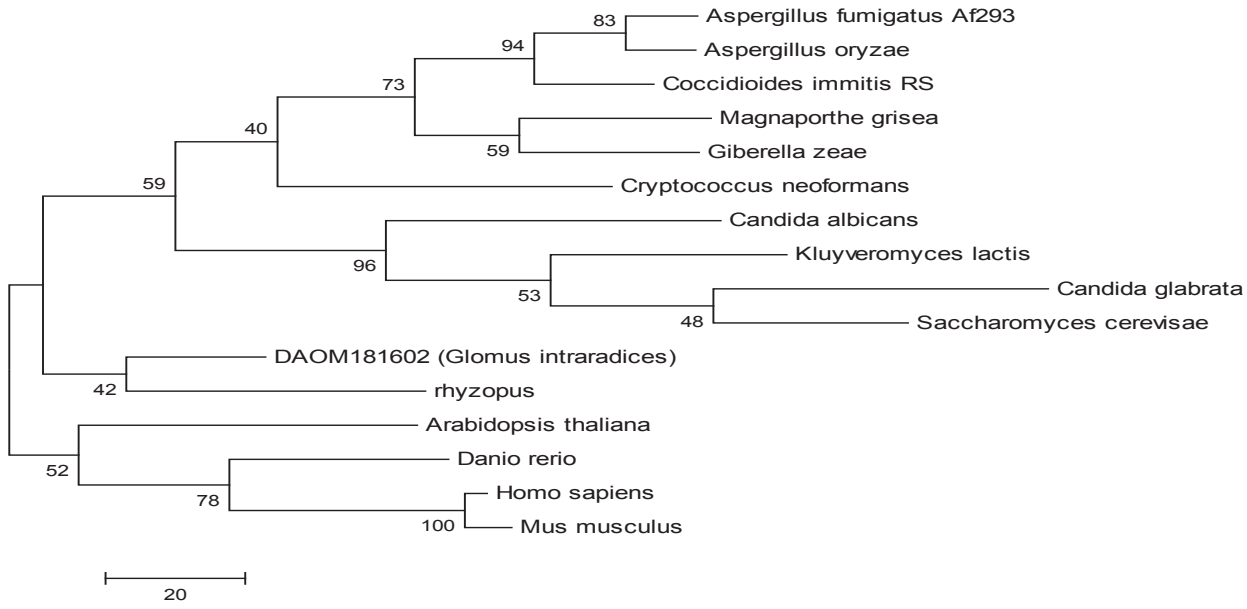


Figure 2: Phylogenetic analysis of meiosis specific msh4 gene (1230 amino acid length). Glomeromycota msh4 sequence was compared with previous published sequences from fungi, mammals, fishes, protist and plants. Numbers at nodes correspond to bootstrap support for neighbor-joining analyses. Scale bar represents 20 substitutions per site.

Numerous studies describe dmc1 in detail and thus we know which nucleotides are important and which domains have to be present in order the protein to be functional. (Proudfoot *et al.*, 2006) These motifs are present in *G. intraradices*, as indicated in Figure 3.

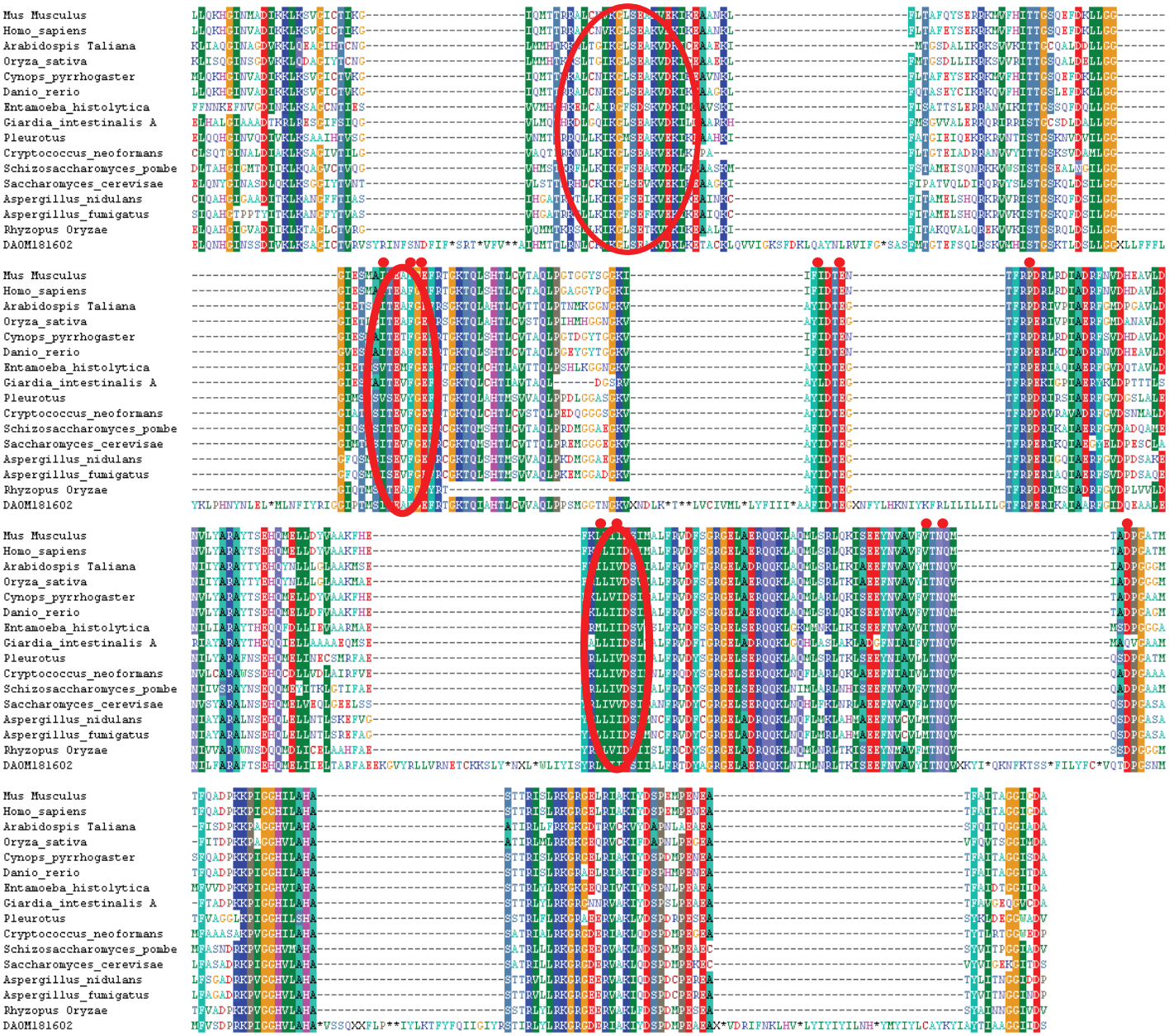


Figure 3: Alignment of *G.intraradices* dmc1 amino acid sequence (DAOM 181602) with previous published sequences from fungi, mammals, fishes, protist and plants. Important domains and amino acids are indicated with red circles and red dots. The Walker A and Walker B motif are present as well as catalytic and structural residues.

Expression data

Different transcripts were found for the msh4 gene. The introns were not always removed. It is probably some intermediate messenger and not alternative splicing because of the presence of STOP codons in each intron. Surprisingly, cDNA found for dmc1 was totally unspliced (Figure 4). Out of a number of clones, I never found a transcript without introns.

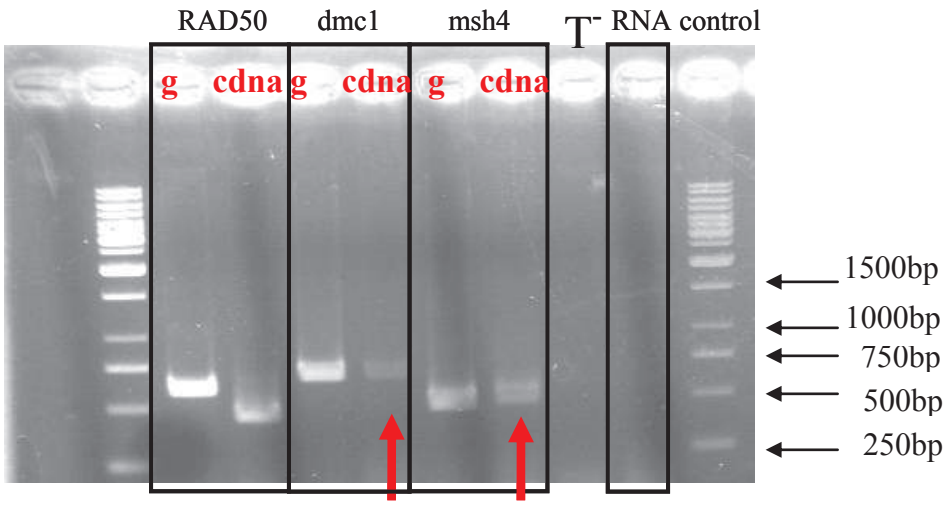


Figure 4: Amplification of RAD50, dmc1 and msh4. For each gene, we amplified from genomic DNA template (g) and cDNA template (cdna). cDNA contaminations were controlled by water amplification (T-) and by a RNA amplification (RNA control). Red Arrows show transcripts of the meiosis specific genes.

Discussion

Using different approaches, we found several genes in *G. intraradices* that could play a role in recombination and meiosis. These genes are shown in the Table 1. Ramesh *et al.* (2005) defined 7 genes as "Core meiotic drive". These 7 genes are potentially sufficient for the meiosis process. These genes are spo11, dmc1, msh4, msh5, Rad50, Rad51 and mlh1. They are all found in *G. intraradices* except spo11. Spo11 initiates meiosis by creating a double strand break (Table1).

Purifying selection is exerted on genes that are used and are essential to the organism. Those which are no longer functional rapidly accumulate mutations and become pseudogenes. If AMF do not undergo meiosis then the presence of conserved meiosis-specific genes would not be expected.

The perfect conservation of these domains suggests that dmc1 genes present in *G. intraradices* is putatively functional (Figure 3). Moreover, the expression of dmc1, but also msh4, supports the fact that those genes have some functional role in the fungus. There are two possibilities to explain why no splice form of dmc1 was observed. The first is that dmc1 is not functional anymore. The second could be that the low level of expression of this gene could explain that we have only the primary transcript (Figure 4).

Taking into account the expression and the conservation of the amino acid sequence, two hypotheses could be advanced:

The first is that meiosis exists in this organism, even if no visual evidence is present. There are several examples, and especially in fungi, where species have been considered asexual until someone detected cryptic sexual eg in *Aspergillus fumigatus* or *Coccidioides immitis* (Burt *et al.*, 1996; Dyer *et al.*, 2005; Mikheyev *et al.*, 2006; O'Gorman *et al.*, 2009). However, the nuclei in *G.intraradices* are haploid (Hijri *et al.*, 2004) and meiosis would need to duplicate its genomic material. Another possibility is that nuclei fuse in order to allow homologous chromosomes to come together. To date, there is no observation of any nuclear fusion in AMF.

An alternative explanation could be that these genes are still under strong selection because they have another function in another pathway. It could explain why no one has observed meiosis in AMF and explain why these genes are so conserved in *G. intraradices*, even if no others functions for these genes have been observed in the entire kingdom of life.

However, since starting this work, the main hypothesis is now strongly supported by the recent study of Croll *et al.* (2009), who showed that recombination had occurred in *G. intraradices* population. Moreover, even if some essential genes like spo11 have not yet been found not in *G. intraradices*, some organisms, such as *Drosophila melanogaster* do not have this gene but are able to undergo meiosis.

Complementary tests on yeast could give us an idea about their function and a western-blot followed by immune-precipitation could allow us to find an interacting protein. If the same interacting proteins are observed as those observed in meiosis and the cellular localization is the same too, it would be supplementary evidence supporting the existence in *G. intraradices*.

Annexe IV

Segregation in a mycorrhizal fungus alters rice growth and symbiosis-specific gene transcription

Caroline Angelard, Alexandre Colard, Hélène Niculita-Hirzel, Daniel Croll, Ian R. Sanders.

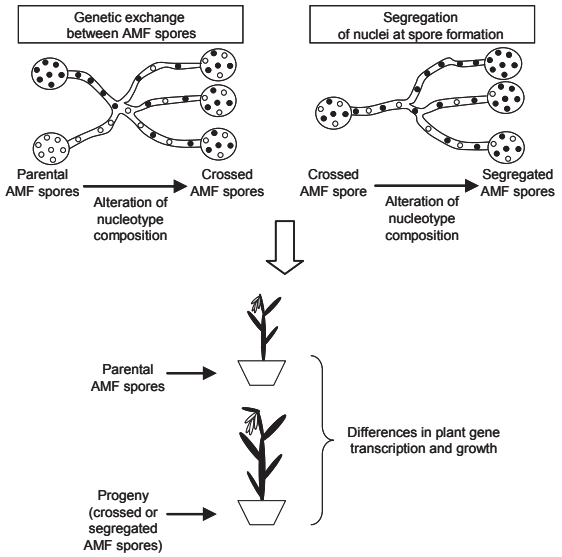
Department of Ecology and Evolution, University of Lausanne, Biophore, Lausanne, CH.

This chapter was published 2010 in *Current Biology*

Abstract

Arbuscular mycorrhizal fungi (AMF) form symbioses with the majority of plants, improving plant nutrition and promoting plant diversity (Harrison, 1997; van der Heijden *et al.*, 1998). Evidence exists suggesting that AMF contain populations of genetically different nucleotypes coexisting in a common cytoplasm (Kuhn *et al.*, 2001; Hijri *et al.*, 2005. This potentially has two important consequences for their genetics. Firstly, by random distribution of nuclei at spore formation, the new clonal offspring of an AMF could receive different complements of nucleotypes compared to the parent or other siblings - we consider this as segregation. Secondly, genetic exchange between AMF would allow the mixing of nuclei, thereby altering the genetic diversity of nucleotypes in newly formed spores. Because segregation was assumed not to occur (Pawlowska *et al.*, 2004; Bever *et al.*, 2005; Pawlowska *et al.*, 2005) and genetic exchange has only recently been demonstrated (Croll *et al.*, 2009, no attempts have been made to test whether this affects the symbiosis with plants. Here we show that segregation occurs in the AMF species *Glomus intraradices* and can enhance the growth of rice up to five times, even though neither parental nor crossed AMF lines induced a positive growth response. We also found that this process results in an alteration of symbiosis-specific gene transcription in rice. Our results demonstrate that simple manipulations of the genetics of AMF have important consequences for their symbiotic effects on plants and could potentially be used to enhance the growth of globally important crops.

Graphical abstract:



Highlights

- Segregation in mycorrhizal fungi revealed by genetic differences among single spore progeny
- Genetic exchange and segregation in mycorrhizal fungi differentially alter plant growth
- Segregation in mycorrhizal fungi can greatly enhance rice growth
- Segregation in mycorrhizal fungi alters symbiosis specific gene transcription in rice

Results and discussion

Testing for the occurrence of segregation in AMF

Despite the enormous importance of the mycorrhizal symbiosis, no attempts have been made to manipulate the genetics of the fungal partners in order to alter their symbiotic effects on plant growth and nutrition. These putative ancient asexuals are thought to grow clonally (Remy, 1994; Judson *et al.*, 1996; James *et al.*, 2006). Because of this it was thought that genetic exchange and segregation were lacking in these fungi. Additionally, little evidence exists for strong positive symbiotic effects on the growth of some of the most globally important crop plants, such as rice. Our aim, in this study, was to start with AMF lines of the species *Glomus intraradices* that do not display positive growth effects on rice and test whether we could induce genetic change, giving rise to genetically novel AMF that would induce different growth effects. A further aim was to test whether these novel AMF lines could also alter the transcription of symbiosis-specific genes in rice.

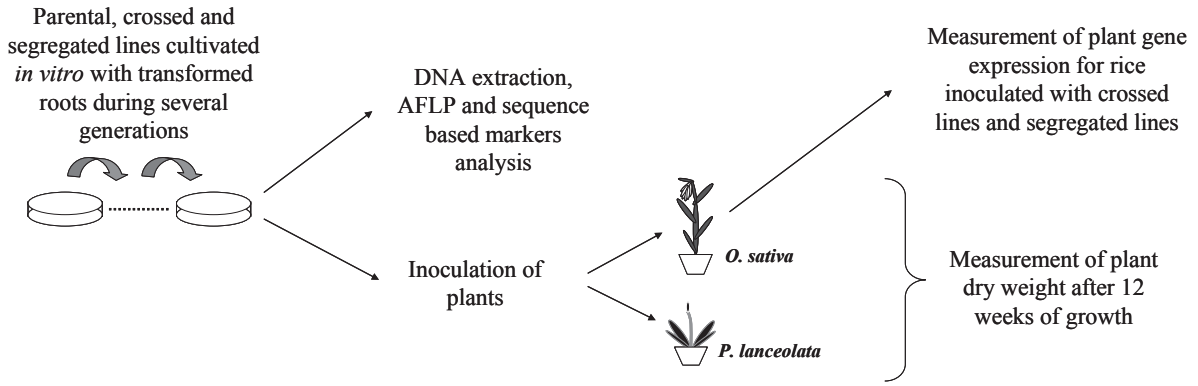


Figure 2.1 Experimental design.

Segregation in AMF has not previously been demonstrated. The likelihood of segregation occurring is dependant on the genetic diversity of nucleotypes present in the mother spore. Genetic exchange should lead to progeny (hereby called crossed lines) with higher genetic diversity compared to their parents because of the mixing of nuclei. Thus, to test the occurrence of segregation, we initiated single spore cultures from crossed lines of the AMF *G. intraradices* (Croll *et al.*, 2009 and Figure 2.1). The number of nuclei present per AMF spore can be from approximately 800- to several thousand, depending on the species (Hosny *et al.*, 1998). The number of nuclei in *G. intraradices* spores has not been measured precisely but is certainly at least several hundred per spore. These crossed lines were previously shown to

exhibit bi-parental inheritance and different phenotypes compared to their parents (Croll *et al.*, 2009. The parental lines originally used to obtain crossed lines were shown to differentially affect plant growth (Koch *et al.*, 2006). Four different crossed lines were used (S3 and S4 from the pairing between parental lines C2 and C3, Sc1 and Sc2 from the pairing between parental lines C3 and D1). Single spores were separately cultivated from each crossed line, giving three to six potentially segregated lines (S3a-f, S4a-c, Sc1a-f and Sc2a-f for crossed lines S3, S4, Sc1 and Sc2, respectively). A previous study has tested the segregation hypothesis in *G. etunicatum* by looking at segregation of 13 different alleles of one locus (Pawlowska *et al.*, 2004). That study did not reveal total segregation of alleles. However, the different alleles could have been shared by many nuclei so that total segregation of any of the variants would not have been detected. Moreover, the study did not look at differences in relative frequencies of the different alleles (partial segregation). Increasing the number of markers would increase the chance to observe segregation that may involve rare nuclei. In the present study, we used both AFLP, as it represents a consensus of the global diversity of one AMF, and sequence based markers. In addition to indicating the presence or absence of markers to reveal total segregation, those methods can also be used to detect changes in nucleotype frequency. Indeed, AFLP can detect partial segregation if some alleles increase in frequency in progeny and become detectable by AFLP, or on the contrary, decrease in frequency and pass below the level of detection by AFLP. Sequence based markers, that allow the amplification of multiple alleles present in a single individual, can also be used to detect quantitative differences by studying changes in the relative frequencies of the different alleles (see Supplemental Experimental Procedures).

We firstly analysed AFLP fragments of potentially segregated lines for the presence and absence of markers that were previously found in the crossed lines (Table 2.1, Table 2.2 and see Tables S1 and S2). Our results show that segregation occurred because reproducible polymorphisms were observed between crossed lines and segregated lines and among segregated lines (Table 2.1 and Table 2.2). See Table S1 for an example of the scoring of the polymorphic markers. The majority of the genetic differences between crossed and segregated lines (52% - 82%) corresponded to the disappearance of markers in at least one of the segregated lines (Table 2.1). A lower percentage of changes (13% - 43%) corresponds to markers that were detected in at least one segregated line and undetected in crossed lines, but which were detected in the original parental lines used to produce the crossed lines (Table 2.1). Finally, 5% -10% of markers were only detected in the segregated lines (Table 2.1).

Disappearance of markers could mean that some nucleotypes were completely segregated among the spores (total segregation). An alternative hypothesis is that the disappearance or appearance of markers could be due to new mutations arising or recombination. Such results could also be due to changes in the frequency of the nucleotypes in the new segregated lines (partial segregation), allowing or preventing their detection by AFLP. The change in nucleotype frequencies is strongly supported by the fact that some markers were detected in parental lines, not detected in crossed lines and then detected again in some segregated lines, suggesting that such markers were at undetectable levels but still present in the crossed lines. Finally, we also found that segregated lines originating from the same crossed line were significantly different from each other when comparing the identity of the polymorphic markers (Table 2.2).

Table 2.1. Number of AFLP markers differing in at least one segregated line and originating from the same crossed line.

Comparison of markers between crossed lines and segregated lines	Sc1 and Sc1a-f	Sc2 and Sc2a-f	S3 and S3a-f	S4 and S4a-c
	Number of markers disappearing in at least one segregated line compared to crossed lines			
Markers detected in parental lines	75	42	54	38
Markers undetected in parental lines	3	6	6	8
Total	78	48	60	46
	Number of markers appearing in at least one segregated line compared to crossed lines			
Markers detected in parental lines	12	27	24	37
Markers undetected in parental lines	5	8	4	4
Total	17	35	28	41

Amplified fragment lenght polymorphism (AFLP) markers markers either were detected in crossed lines and disappeared in at least one segregated line (upper section of table) or were undetected in crossed lines and appeared in at least one segregated line (lower section of table). Polymorphic markers were detected in at least one of the parental lines used to produce crossed lines or were undetected in parental lines. The total number of markers analyzed in crossed lines were 938, 917, 938 and 950 for crossed lines Sc1, Sc2, S3 and S4, respectively. See Table S1 for an example of determination of polymorphic markers and Table S2 for more details.

Sequence based markers were also used to detect partial segregation (or alterations in nucleotype frequency). The locus Bg112 presents multiple alleles within the same individual and the alleles differ in sequence and in the number of tri-nucleotide repeats. We first used a pair of primers flanking the repeat region to amplify alleles that could be discriminated by

sequence and length. Qualitative differences among AMF lines do not occur because all the alleles were detected in each AMF line (Figure 2.2A). However, we then compared the relative frequencies of the different alleles among AMF lines to test the occurrence of partial segregation (see Supplemental Experimental Procedures and Table S3). MANOVA results revealed strong significant differences in the proportion of the four alleles among AMF lines initiated from the two crossed lines S3 and Sc2. For example, the proportion of allele 1 decreased in the segregated lines produced from the crossed line S3, while the proportion of allele 4 either remained stable or increased in these segregated lines (Figure 2.2A). The proportions of the four alleles did not differ significantly among the AMF lines initiated from crossed lines Sc1 and S4 (Figure 2.2A). We also used another pair of primers down-stream of the repeat region to amplify alleles from the crossed line S3 and the segregated lines S3d, S3e and S3f that could be distinguished by sequence. Frequencies of the alleles were determined by 454 pyrosequencing of the amplicons. There were considerable differences in the frequency of alleles I- III among S3d, S3e and S3f (Figure 2.2B and Table S3). Furthermore, qualitative differences were also detected among the AMF lines (Figure 2.2B). Allele IV was detected in one of the parental lines used to produce the crossed line S3 (data not shown) and in two segregated lines, but not in the crossed line S3 and in the other segregated line (Figure 2B). Another allele (allele V) was only detected in two of the segregated lines (Figure 2.2B).

The disappearance or the fluctuation in frequency of nucleotypes shown in the present study by AFLP and sequence based markers means that new spores formed from the same individual can contain a different assortment or a different proportion of nucleotypes compared to parental or other new spores. Thus, segregation (total or partial) occurred in *G. intraradices*.

Table 2.2. Number of markers differing in each segregated line compared to the crossed line of origin.

Segregated Line	Nb of markers disappearing in segregated line	Nb of markers appearing in segregated line	G_{adj}	ddl	p value
S3a	30	15			
S3b	14	13			
S3c	30	7			
S3d	13	11			
S3e	30	8			
S3f	28	9			
Proportion of polymorphic markers			10.62	5	0.059
Proportion of disappearing and appearing markers			11.04	5	0.051
S4a	26	31			
S4b	25	37			
S4c	23	3			
Proportion of polymorphic markers			17.05	2	<0.001
Proportion of disappearing and appearing markers			19.84	2	<0.001
Sc1a	57	3			
Sc1b	30	5			
Sc1c	22	5			
Sc1d	23	2			
Sc1e	28	3			
Sc1f	14	11			
Proportion of polymorphic markers			24.72	5	<0.001
Proportion of disappearing and appearing markers			19.28	5	0.002
Sc2a	23	9			
Sc2b	19	7			
Sc2c	11	25			
Sc2d	18	19			
Sc2e	9	13			
Sc2f	13	16			
Proportion of polymorphic markers			5.69	5	0.34
Proportion of disappearing and appearing markers			18.30	5	0.003

The rightmost three columns correspond to the results of the G tests used to detect significant genetic differences among arbuscular mycorrhizal fungus lines originating from crossed lines: adjusted G value (G_{adj}), degrees of freedom (df), and significance (p value).

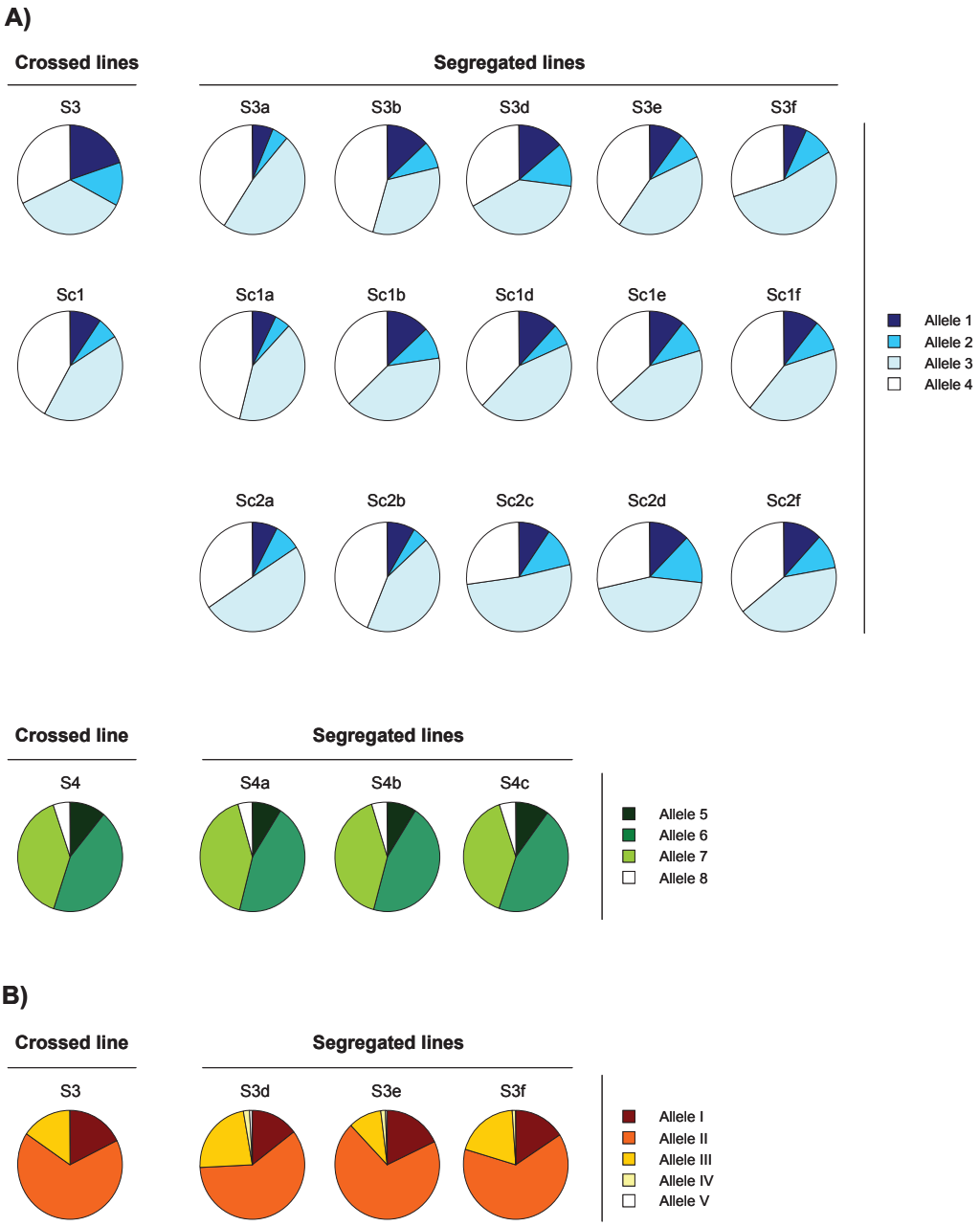


Figure 2.2 Relative allele frequencies of the locus Bg112 in crossed and segregated AMF lines. A) Alleles differed in their length and the allele frequencies were measured as relative fluorescence units (RFU) from electrophenograms. The same four alleles were found for the crossed lines S3, Sc1 and Sc2 and their segregated lines (Alleles 1-4) and four other alleles were found for the crossed line S4 and the segregated lines S4a-c (Alleles 5-8). The amount of DNA was insufficient for the crossed line Sc2 and the segregated lines S3c, Sc1c and Sc2d to perform the analyses. MANOVA results: Segregation in line S3: Pillai's Trace = 1.48, approximated $F_{20,80}$ = 2.36, p = 0.004; Segregation in line Sc1: Pillai's Trace = 0.86, approximated $F_{20,80}$ = 1.10, p = 0.4; Segregation in line Sc2: Pillai's Trace = 1.40, approximated $F_{16,68}$ = 2.30, p = 0.009; Segregation in line S4: Pillai's Trace = 0.62, approximated $F_{12,39}$ = 0.85, p = 0.6. B) Allele frequencies differed according to 454 pyrosequencing on DNA of the crossed line S3 and the segregated lines S3d, S3e and S3f. Five alleles were found (Alleles I-V). Appearance and disappearance of alleles was also observed. See also Table S3 for numeric values.

Effects of genetic exchange and segregation on plant growth

Having detected segregation in AMF, we then tested the effects of both genetic exchange and segregation in AMF on plant growth. We inoculated two plant species (*Plantago lanceolata* and *Oryza sativa*) with parental, crossed and segregated lines (Figure 2.1). *P. lanceolata* is a plant that has repeatedly been shown to benefit from mycorrhizal fungi, while few, if any reports clearly demonstrate strong benefit of AMF inoculation on rice growth. The effect of genetic exchange in AMF on plant growth was tested by inoculating both plant species with six crossed lines and their respective parental lines. Genetic exchange in AMF significantly altered plant growth (Figure 2.3A and Table S4). Where significant changes occurred, the growth of both plant species was reduced after inoculation with crossed lines compared to plants inoculated with parental lines (Figure 2.3A). Moreover, *O. sativa* inoculated with either crossed or parental lines did not grow any larger than uninoculated plants (Figure 2.3A). We then tested the effect of segregation in AMF on plant growth. Plants were inoculated with three crossed lines and their respective segregated lines. In all cases, except for *P. lanceolata* inoculated with crossed line S3 and the subsequent segregated lines, segregation in AMF had large effects on plant growth (Figure 2.3B and Table S4). In *P. lanceolata*, when significant differences occurred, we found that segregated lines reduced plant growth compared to inoculation with crossed lines (Figure 2.3B). This was the case for the segregated line Sc2b compared to crossed line Sc2 and segregated line Sc4 compared to crossed line S4 (Figure 2.2B). However, we found the opposite in rice. Indeed, *O. sativa* inoculated with some segregated lines (Sc2d and Sc2e; Figure 2.3B) were 2-5 times larger than plants inoculated with other AMF lines, including the crossed line, and also compared to uninoculated plants (Figure 2.3B). One offspring line of S3 and one offspring line of S4 also resulted in significantly higher rice growth than plants inoculated with S3 or S4 (Figure 2.3B). The differences in how the two plant species respond to inoculation with segregated lines are highlighted by the significant AMF line by plant species interactions (Table S4). In all cases where significant differences in plant growth (either positive or negative) occurred after inoculation with a segregated AMF line, compared to inoculation with the crossed line of origin, genetic differences had been demonstrated for those lines either with AFLP or frequency of alleles at the Bg112 locus, or by both methods. Our work shows that genetic changes in the fungus affect plant growth, but further long-term studies need to address how stable these effects are over time.

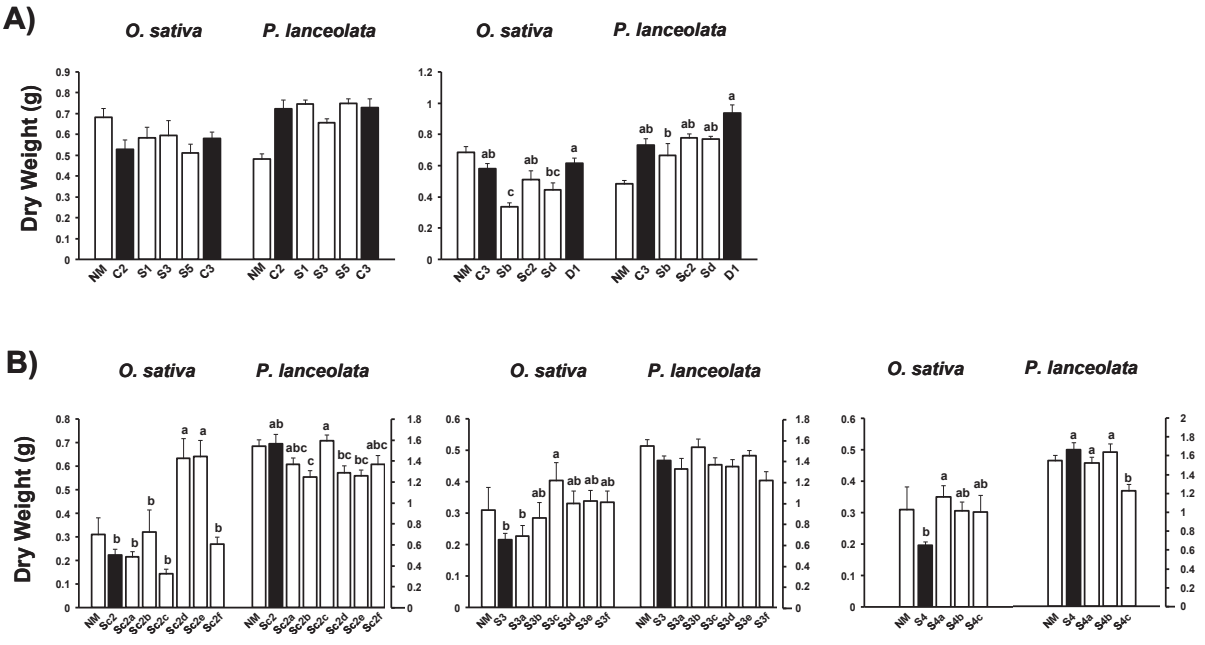


Figure 2.3 Effect of genetic exchange and segregation in AMF on plant growth. A) Mean dry weight of non-mycorrhizal (NM) plants (dashed columns), plants inoculated with parental lines (black columns) and crossed lines (white columns). S1, S3 and S5 come from pairing C2xC3, and Sb, Sc2 and Sd come from pairing C3xD1. The overall effect of AMF inoculation was plant species specific: *P. lanceolata* significantly benefited from being mycorrhizal ($F_{1,90}$=33.73, $p < 0.001$) and inoculation significantly reduced growth in rice ($F_{1,94}$=9.20, $p < 0.01$). B) Mean dry weight of NM plants (dashed columns), plants inoculated with crossed lines (black columns) and segregated lines (white columns). No overall effect of AMF inoculation was found ($F_{1,180}$=2.78, $p = 0.1$ and $F_{1,180}$=0.24, $p = 0.62$ for *P. lanceolata* and the rice, respectively) because of the contrasting effects on plant growth of the AMF lines. Significant differences between NM and mycorrhizal plants: *O. sativa* inoculated with Sc2d and Sc2e; *P. lanceolata* inoculated with Sc2b, Sc2e, S3f and S4c, according to the Dunnett's test ($p < 0.05$). Error bars represent + 1 SE and different letters above columns indicate a significant difference ($p < 0.05$) according to Tukey-Kramer HSD test. See also Table S4.

Even though rice establishes symbioses with AMF, the present study shows that inoculation with non-manipulated AMF does not induce a growth benefit. This is probably why there are no reports in the literature showing that rice benefits from AMF inoculation, despite the vast wealth of published studies showing responses of hundreds of plant species to inoculation with AMF. However, our results clearly demonstrate that such responses are not fixed and can be altered by simple manipulations of the genetics of the fungi by genetic exchange and segregation. Additionally, the effects found in both experiments were plant species specific and, thus, the effects of genetic changes induced in AMF were not the same on the symbiosis with two different plant species. Other studies found that plants did not respond in the same

way when inoculated with different AMF (Johnson *et al.*, 1997; StreitwolfEngel *et al.*, 1997; van der Heijden *et al.*, 1998; Koch *et al.*, 2006), and that AMF are not always beneficial for all plant species (Johnson *et al.*, 1997; Klironomos *et al.*, 2003; Koch *et al.*, 2006). However, those studies only considered the effects of AMF diversity at the community and population levels. Our results demonstrate that manipulating the genetics of AMF at the individual level has important consequences for symbiosis with plants, even those of global importance for human nutrition.

Effects of segregation on symbiosis-specific gene transcription

We looked at the effects of segregation in AMF on symbiosis specific gene transcription. Rice genes that are specifically up-regulated during the establishment of the symbiosis have recently been identified (Paszkowski *et al.*, 2002; Guimil *et al.*, 2005; Gutjahr *et al.*, 2008). We analysed the expression of four of these genes: AM1, AM3, AM14 and PT11 in rice. Two of the genes (AM1, a putative peroxidase and AM3 that contains a peptidoglycan binding LysM domain 1) are genes transcribed early in rice following colonization by the fungus. This means that their transcription begins even before the formation of arbuscules (the fungal structures inside root cells allowing exchange of nutrients between plants and AMF). The other two genes (AM14, a serine-threonine kinase-like gene and PT11, a phosphate transporter) are transcribed later in the establishment of the symbiosis (following arbuscule formation). PT11 is strictly induced in arbusculated cells. This important gene is implicated in transfer of phosphorous, the major nutrient transported by the fungus in the AM symbiosis. The effect of segregation in AMF on the transcription of these genes was studied in rice inoculated with the crossed line S4 and the segregated lines S4a-c. In the segregation experiment, the lines S4a-c did not result in the largest differences in rice growth, but due to the length of time taken to carry out such experiments, the rice gene experiment had to be initiated before the plant growth results were available. Results show that segregation in AMF significantly affected gene transcription in rice (Figure 2.4). The differential effects of segregated AMF lines on gene transcription were similar in the four genes studied. Inoculation with segregated lines S4a and S4c induced higher transcription of the four genes compared to plants inoculated with S4b or crossed line S4. Although the patterns of transcription of all four genes seem to be correlated, this was not the case for genes not involved in symbiosis. Transcription of two housekeeping genes, encoding polyubiquitin and cyclophilin were also measured. Because transcription of the polyubiquitin gene did not differ among plants inoculated with different AMF lines, this was used to standardize the data.

Secondly, transcript levels of the housekeeping gene encoding cyclophilin differed significantly among plants inoculated with different AMF lines but they were not correlated with the patterns of the symbiosis-specific gene transcription (see Figure S1). Thus, we show that although transcription levels of the 4 symbiosis-specific genes were correlated with each other, they are independent of transcription levels of some other rice genes that are not involved in the symbiosis. Our results show that segregation in AMF can have large effects, not only in genes involved in plant growth, but also, on early and late transcribed genes that are specific to the establishment of the symbiosis, and on fundamental mycorrhizal specific

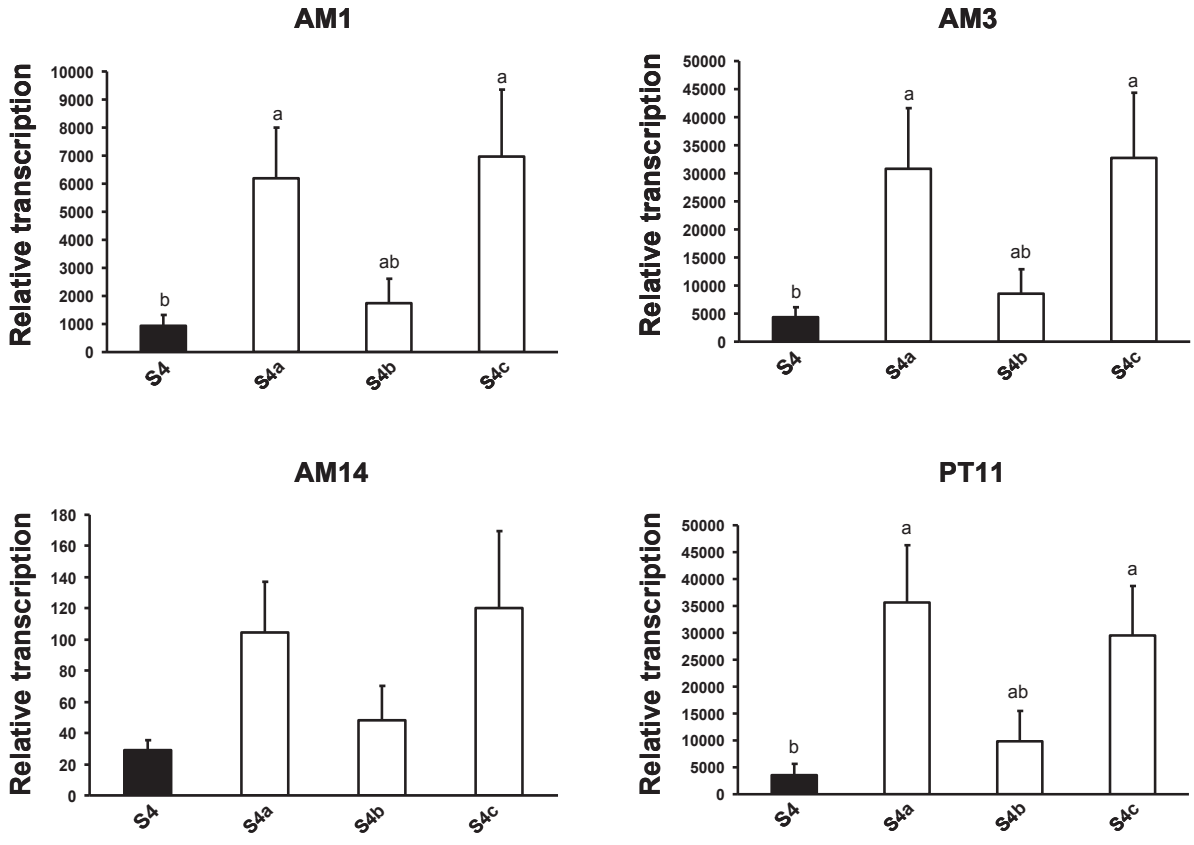


Figure 2.4 Effect of segregation in AMF on plant gene transcription. Real-time RT-PCR-based transcription analysis of AM1, AM3, AM14 and PT11 genes of rice roots inoculated with crossed line S4 (black bar) and segregated lines S4a-c (white bars). Transcription levels are shown relative to the constitutively expressed *polyubiquitin* gene. Error bars represent + 1 SE and different letters above bars indicate a significant difference ($p < 0.05$) according to Tukey-Kramer HSD test. See also Table S5 and Figure S1.

genes implicated in nutrient acquisition in plants. Further studies adopting a microarray or high-throughput sequencing approach are highly warranted as they could give a wider understanding of how genetic changes in the fungus affect overall patterns of plant gene expression, extending beyond known symbiosis-specific genes.

Genetic exchange and segregation in AMF are two processes that can potentially alter the nucleotype complement of new AMF progeny. If genetically different nucleotypes express different genes, then alteration in nucleotype frequencies could have important consequences for the symbiosis. Overall, our results demonstrate that genetic exchange and segregation in AMF can significantly affect how the plant responds to the symbiosis, resulting in species-specific negative or positive growth effects and altering the transcription of plant genes that are specifically up-regulated during the establishment of the symbiosis. The exact mechanisms of how changes in nucleotype frequency in the fungus alter the symbiotic effects on plants still remains a "black box" and future research efforts should be directed at unravelling such mechanisms. Our results represent an important step in the understanding of AMF genetics and form a basis for further experimental approaches to understand how AMF genetics can be used to promote the growth of globally important crops.

Experimental procedures

Detailed information on methods is provided in the Supplemental Experimental Procedures. Briefly, all the AMF lines used in the experiments were cultivated on plates containing a non-mycorrhizal Ri T-DNA-transformed carrot roots (Becard *et al.*, 1988) and maintained through clonal subculturing (growth period of 15 weeks) with same protocol as in Koch *et al.* (2004). The hyphae and spores produced clonally from each parental, crossed and segregated line were used to perform two independent DNA extractions and subsequent molecular analysis. AFLP and sequenced based markers were used to detect segregation (total or partial) in AMF. To test the effect of genetic exchange and segregation on plant growth, the dry mass (root plus shoot) of each plant was measured after a growth period of 12 weeks. To test the effect of segregation on symbiosis-specific gene transcription, analyses were performed on fresh rice roots after the same period.

Acknowledgements

This work was supported by a grant from the Swiss National Science Foundation (grant numbers 31000AO-105790/1 and 31003A-127371). We thank Caroline Gutjahr for help in the quantification of gene transcripts in rice, Uta Paszkowski for providing rice seeds, Jérémy Bonvin for help with DNA extraction and molecular analyses and three anonymous reviewers for helpful comments. H.NH. performed the determination of allelic frequency by pyrosequencing, A.C. performed the quantification of gene transcripts, D.C. provided crossed AMF lines. C.A. designed the study, performed the AFLP analyses, determined allelic frequency by RFU, performed the greenhouses experiments and analysed all the data. C.A. and I.S. wrote the paper. All authors discussed the results and commented on the manuscript.

Supplemental Information

Inventory of supplemental information: The supplemental items include five Supplemental Tables, one Supplemental Figure, Supplemental Experimental Procedures and Supplemental References.

Table S1: the table helps to understand the procedure to determine the identity of the polymorphic markers showed in Table 2.1.
Table S2: the table provides additional data about the identity of the markers of parental and crossed lines presented in Table 2.1.
Table S3: the table provides numeric values of the data presented in Figure 2.2.
Table S4: the table provides the statistical analyses of the data presented in Figure 2.3.
Table S5: the table provides the sequences of the primers used for the quantification of gene transcription by real-time RT-PCR presented in Figure 2.4.
Figure S1: the figure provides additional transcription analysis related to the Figure 2.4.

Supplemental Data

Table S1: Example of the determination of the identity of the polymorphic markers found between at least one segregated line and the crossed line S4, for data shown in Table 2.1.

Loci (bp)	Parental lines		Crossed line	Segregated lines			Identity of the polymorphic markers between at least one segregated line and the crossed line S4	
	C2	C3	S4	S4a	S4b	S4c		
77	1	0	**1**	1	0	0	Markers were detected in at least one of the parental lines used to produce crossed line	Disappearing markers: markers detected in crossed line and undetected in at least one segregated line
83	0	1	**1**	1	1	0		
95	1	1	**1**	0	0	1		
105	0	0	**1**	1	0	0	Markers were undetected in parental lines	
108	1	0	**0**	1	0	0	Markers were detected in at least one of the parental lines used to produce crossed line	Appearing markers: markers undetected in crossed line and detected in at least one segregated line
120	0	1	**0**	1	1	0		
125	1	1	**0**	0	0	1		
129	0	0	**0**	1	0	0	Markers were undetected in parental lines	

Presence (1) or absence (0) of an AFLP fragment of a given length.

Table S2: Summary of AFLP markers for the parental and crossed lines, for data shown in Table 2.1.

	Number of parental lines specific markers			Markers common to both parental lines	Added markers in crossed lines compared to parental lines	Total number of markers	Markers with an unknown state
Parental lines	**C2**	**C3**					
C2	459	0		489		948	40
C3	0	439		489		928	44
Crossed lines	Presence of parental lines specific markers in crossed lines			Presences of common parental markers			
S3	23	425		476	9	933	54 (45)
S4	429	34		476	11	950	35 (24)
S5	15	416		482	9	922	72 (48)
Parental lines		**C3**	**D1**				
C3		414	0	521		935	44
D1		0	439	521		960	13
Crossed lines		Presence of parental lines specific markers in crossed lines		Presences of common parental markers			
Sb		396	23	509	5	933	57 (48)
Sc1		401	21	513	3	938	65 (44)
Sc2		387	21	503	6	917	69 (54)

Numbers between parentheses correspond to the number of markers with an unknown state after removing the markers with an unknown state in the parental lines.

Table S3: Mean ± SE of the relative peak heights used to compare the relative frequency of the different alleles shown in Figure 2.2A and proportion of the different alleles shown in Figure 2.2B.

Alleles	1	2	3	4	
AMF lines					
S3	0.80±0.16	0.51±0.10	1.41±0.16	1.29±0.13	
S3a	0.25±0.06	0.20±0.05	1.91±0.09	1.64±0.02	
S3b	0.54±0.09	0.32±0.08	1.32±0.04	1.82±0.17	
S3d	0.56±0.10	0.51±0.07	1.60±0.13	1.32±0.10	
S3e	0.41±0.03	0.31±0.06	1.67±0.03	1.61±0.11	
S3f	0.29±0.03	0.37±0.05	2.13±0.14	1.20±0.16	
Sc1	0.39±0.09	0.25±0.06	1.67±0.08	1.68±0.08	
Sc1a	0.31±0.04	0.18±0.03	1.67±0.09	1.84±0.05	
Sc1b	0.53±0.11	0.38±0.07	1.60±0.09	1.49±0.10	
Sc1d	0.49±0.11	0.24±0.08	1.75±0.05	1.51±0.08	
Sc1e	0.43±0.05	0.39±0.04	1.71±0.06	4.47±0.03	
Sc1f	0.43±0.07	0.38±0.10	1.64±0.07	1.56±0.12	
Sc2a	0.31±0.06	0.32±0.06	1.99±0.08	1.38±0.05	
Sc2b	0.35±0.02	0.19±0.02	1.71±0.10	1.75±0.11	
Sc2c	0.39±0.01	0.46±0.04	2.05±0.02	1.09±0.03	
Sc2d	0.50±0.09	0.58±0.07	1.79±0.14	1.14±0.02	
Sc2f	0.48±0.02	0.42±0.01	1.67±0.05	1.44±0.03	
Alleles	5	6	7	8	
AMF lines					
S4	0.44±0.02	1.76±0.02	1.58±0.02	0.22±0.01	
S4a	0.37±0.01	1.79±0.02	1.67±0.02	0.18±0.01	
S4b	0.37±0.01	1.80±0.02	1.64±0.01	0.19±0.01	
S4c	0.42±0.02	1.80±0.01	1.59±0.03	0.23±0.01	
Alleles	I	II	III	IV	V
AMF lines					
S3	0.18	0.67	0.15	0	0
S3d	0.15	0.59	0.23	0.02	0.01
S3e	0.18	0.70	0.10	0.01	0.01
S3f	0.16	0.63	0.20	0.01	0

Table S4: Analysis of variance on the dry weight of plants in the genetic exchange and segregation experiments, for data shown in Figure 2.3.

	AMF lines		Plants		AMF lines x Plants	
Genetic exchange experiment						
Pairing C2 x C3	$F_{4,85}$=0.39	ns	$F_{1,85}$=37.51	***	$F_{4,85}$=1.27	ns
Pairing C3 x D1	$F_{4,82}$=8.45	***	$F_{1,82}$=90.55	***	$F_{4,82}$=1.41	ns
Segregation experiment						
Segregation on line Sc2	$F_{6,120}$=2.50	**	$F_{1,120}$=894.56	***	$F_{6,120}$=12.09	***
Segregation on line S3	$F_{6,121}$=1.67	ns	$F_{1,121}$=1065.37	***	$F_{6,121}$=2.18	*
Segreagtion on line S4	$F_{3,67}$=4.97	**	$F_{1,67}$=856.89	***	$F_{3,67}$=7.93	***

Significance level are: ***, $p < 0.001$; *, $p < 0.05$; ns, $p > 0.05$

Table S5: Primers used for quantification of gene transcription by real-time RT-PCR, for data shown in Figure 2.4.

Gene	Primer Sequences 5' → 3'	
*Os*PT11	Forward:	GAGAAGTTCCCTGCTTCAAGCA
	Reverse:	CATATCCCAGATGAGCGTATCATG
*Os*AM1	Forward:	ACCTCGCCAAAATATATGTATGCTATT
	Reverse:	TTTGCTTGCCACACGTTTTAA
*Os*AM3	Forward:	CTGTTGTTACATCTACGAATAAGGAGAAG
	Reverse:	CAACTCTGGCCGGCAAGT
*Os*AM14	Forward:	CCAACACCGTTGCAAGTACAATAC
	Reverse:	GCACTTTGAAATTGGACTGTAAGAAA
*Os*Ubiquitin	Forward:	CATGGAGCTGCTGCTGTTCTAG
	Reverse:	CAGACAACCATAGCTCCATTGG
*Os*Cyclophilin2	Forward:	AGCTCTCCTAGATCTGTGCTG
	Reverse:	GCGATATCATAGAACGAGCGA

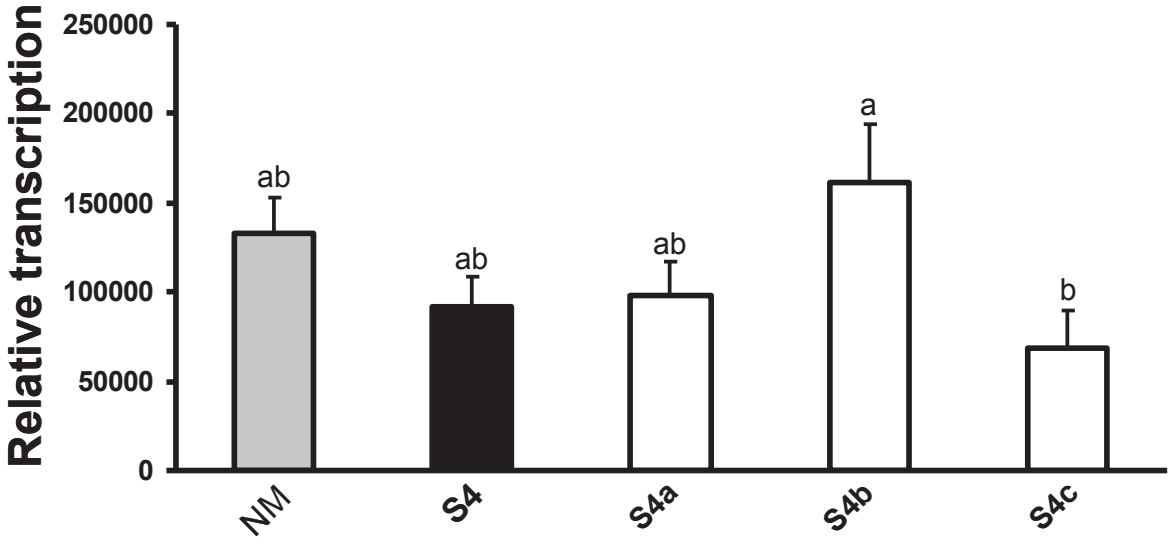


Figure S1: Transcription analysis of Cyclophilin gene, related to Figure 2.4. Real-time RT-PCR-based transcription analysis of rice roots not inoculated with AMF (NM, grey bar), inoculated with crossed line S4 (black bar) and segregated lines S4a-c (white bars). Transcription levels are shown relative to the constitutively expressed *polyubiquitin* gene. Error bars represent + 1 SE and different letters above bars indicate a significant difference ($p < 0.05$) according to Tukey-Kramer HSD test. The statistical results do not change by removing the NM treatment.

Supplemental Experimental Procedures

Cultivation of Glomus intraradices

All the AMF lines used in the experiments were cultivated on plates containing a non-mycorrhizal Ri T-DNA-transformed carrot roots (Becard *et al.*, 1988) and maintained through clonal subculturing (growth period of 15 weeks) with same protocol as in Koch *et al.* (2004). Three parental lines of *G. intraradices* (C2, C3 and D1) were used. These lines were previously shown to be genetically distinct, to differ in their phenotypic traits (Koch *et al.*, 2004; Croll *et al.*, 2009) and to differentially affect plant growth (Koch *et al.*, 2006). In a previous study, two pairs of these parental lines (C2xC3 and C3xD1) were co-cultured together to obtain crossed lines exhibiting bi-parental inheritance (Croll *et al.*, 2009). In the present study, four crossed lines (S1, S3, S4 and S5) coming from the pairing C2xC3 and four crossed lines (Sb, Sc1, Sc2 and Sd) coming from the pairing C3xD1 were used in the different experiments (Sc1 was not included in the previous study of Croll *et al.* (2009) but also showed a bi-parental inheritance (D. Croll, personal communication) and Sc2 in the present study corresponded to Sc in the previous study). To obtain potentially segregated lines, four crossed lines (S3, S4, Sc1 and Sc2) were used and 80 single spores from each crossed line were isolated. The rate of germination of these single spores was 81%, 80%, 91% and 88% for crossed lines S3, S4, Sc1 and Sc2, respectively. Among the germinated spores, 18 to 29 spores of each crossed line were individually transferred to new plates. Three to 9 single spores of each crossed line successfully colonized the transformed carrot roots and gave a sufficient amount of mycelium to be further cultivated. Three to six of these potentially segregated lines were then used for the experiments (S3a-f for crossed line S3; S4a-c for crossed line S4; Sc1a-f for crossed line Sc1 and Sc2a-f for crossed line Sc2). Prior to all the present experiments, all the parental, crossed and segregated lines were cultivated with a minimum of 4 subculturing events. The hyphae and spores produced clonally from each parental, crossed and segregated line were used to perform two independent DNA extractions and subsequent molecular analysis.

Amplified fragment length polymorphism

AFLP was used to score total genetic variation for all parental, crossed (except S1 and Sd) and segregated lines. The protocols to extract DNA, perform and score AFLP were the same as in Croll *et al.* (2009) except for the following. DNA extractions were performed after six

generations of subculturing. We used 7 combinations of selective primers for the amplification of fragments: EcoRI-AA (HEX)/MseI-TT, EcoRI-AA (FAM)/MseI-TG, EcoRI-TT (HEX)/MseI-TT, EcoRI-TA (HEX)/MseI-TT, EcoRI-AA (FAM)/MseI-TA, EcoRI-TT (HEX)/MseI-TA, EcoRI-AAc (FAM)/MseI-TT. Two independent DNA extractions were performed for each AMF line and band presence was scored if bands of both replicate runs exceeded 50 relative fluorescence units (RFU). We used a conservative analysis where all the markers with an unknown state (see Croll *et al.* (2009) for details about unknown states) in the parental lines (either C2, C3 or D1) were not considered to score the crossed and segregated lines. See Table S2 for a summary of the AFLP markers in parental and crossed lines.

Sequence based markers and detection of allele frequencies

Conventional fingerprinting. The Locus Bg112 comprises a tri-nucleotide repeated pattern and the alleles differ both in their length as well as in deletions, insertions and substitution. A couple of primers, Bg112-up and Bg112-d, have been designed around the tri-nucleotide repeat region in order to discriminate by length different Bg112 alleles. Primers were developed using the library BG112-T3 (accession number: CG431793) and program Tandem Repeat Finder Benson. The PCR reactions conducted with Bg112-up (5'-AAGGTCATATCAAATTCTCCGATCC) and Bg112-d (5'-GAAACTGGGAAGTCAGCTCTTGT) were the same as in Croll *et al.* (2009) except that the number of PCR cycles was adapted. Four alleles have been identified with this pair of primers for the crossed lines S3, Sc1 and Sc2 and their subsequent segregated lines (allele 1: 174bp, allele 2: 178bp, allele 3: 208bp and allele 4: 214bp; accession numbers on GenBank: GU930826, GU930824, GU930827, GU930828, respectively). Four other alleles have been identified for the crossed line S4 and the segregated lines S4a, S4b and S4c (allele 5: 160bp, allele 6: 168bp, allele 7: 174bp and allele 8: 180bp; accession numbers on GenBank: GU930825, GU930836, GU930835, GU930834, respectively)). We replicated the PCR reactions two times with 5 ng of DNA. Alleles and their peak heights (relative fluorescence units, RFU) were manually scored using GeneMapper 4.0^{TM} (Applied Biosystems, Inc.). Four values were then obtained for each allele and each AMF line (two PCR reactions on each of the two independent DNA extractions). The AMF lines S3d, Sc1f, Sc2a and S4 gave sufficient fungal material to make three independent extractions with a sufficient amount of DNA to perform analyses; therefore six values were obtained. However, the quantity of DNA

was not sufficient after having performed the AFLP analyses for the AMF lines S3c, Sc1c, Sc2 and Sc2e. As four alleles of the locus BG112 were amplified in each crossed and segregated line, each allele is not independent from the others; therefore, analysis on the height of each allele separately would be unsuitable. Relative peak heights were thus used to compare the relative frequency of the different alleles (for each DNA extraction and each PCR amplification, peak height of each allele was divided by mean peak height of the four alleles). We had carried out two previous tests on the appropriateness of this method for detecting differences in allele frequency. The first was to amplify different amounts of template fungal DNA and see if this changed the peak heights. We also used replicate samples to look at the variation in peak height among samples of the same amount of template DNA. This was found to be highly reliable. A further test was to take known amounts of template and mix them in different amounts to see if this would give predictable and reliable peak heights. This test also indicated that the technique was reliable. Studies on the comparison of allele frequencies in human populations often use pooled DNA samples from many different individuals. Whole pool comparisons have been shown to accurately detect the changes in frequencies of given alleles. Allele relative fluorescence unit (RFU) is used as an indication of their relative frequencies. Such quantitative techniques are also commonly used in many other studies on communities of bacteria, insects, plant and animal breeding (Barcellos *et al.*, 1997; Daniels *et al.*, 1998; Collins *et al.*, 1999; Kraft *et al.*, 1999; Piepho *et al.*, 2000 ; Jansen *et al.*, 2001; Lueders, 2003). In this study, we used RFU values for the four alleles of the locus Bg112 to look for evidence of changes in allele frequencies. We only considered the height of the four alleles and not stutter peaks because they are too small in comparison to the main peak.

Pyrosequencing

Another method that can be used in allele detection and in comparison of allele frequencies is the pyrosequencing of amplicons. This method has been shown to accurately detect the changes in frequencies of given alleles, the number of reads being used as an indication of an allele's relative frequency (Solmone *et al.*, 2009; Afsar *et al.*, 2010; Muggerud *et al.*, 2010). In this study, another pair of primers, Gp366 (5'-**A**xxxxxxTATGGTGTGAATAAGGATTTAGCTT-3') and Gp3 (5'-**B**CTTCTCAACATTCACGTAAATCTTC-3'), have been designed in order to amplify ~350bp and sequence 250bp down-stream of the tri-nucleotide repeated region in Bg112,

using pyrosequencing. Amplicon libraries were established using these primers. The "xxxxxx" corresponds to a 6 bp tag; one for each fungal line. Their sequence has been chosen as recommended by Roche for the tag-encoded 454 GS-FLX amplicon pyrosequencing method. A and B represent the two pyrosequencing primers (GCCTCCCTCGCGCCATCAG and GCCTTGCCAGCCCGCTCAG). The PCR conditions used were 95°C for 2 min, 30 cycles of 30 s at 94°C (denaturation), 50°C for 45 s (annealing) and 72°C for 90 s (extension), followed by 10 min at 72°C. All PCR products were purified using the Qiagen MinElute kit (Qiagen, Switzerland). The amplicon lengths and concentrations were estimated, and an equimolar mix of all amplicon libraries was used for pyrosequencing. Pyrosequencing of the amplicon library (from the Gp366 primer) on the Genome Sequencer FLX 454 System (454 Life Sciences/Roche Applied Biosystems, Nutley, NJ, USA) at Microsynth (Switzerland)) resulted in 10'000 +/- 1'000 reads per sample that satisfied the sequence quality criteria employed (Droege *et al.*, 2008). Tags were extracted from the FLX-generated composite FASTA file into individual sample-specific files based on the tag sequence by the proprietary software of Roche. The alleles present in each amplicon have been identified and their frequency calculated with GS Amplicon Variant Analyzer Software from Roche. The analyses were performed with DNA of the crossed line S3 and the segregated lines S3d, S3e and S3f. In total, five alleles were identified (alleles I-V; accession number on GneBank: GU930829, GU930830, GU930831, GU930832, GU930833, respectively).

Following the previous molecular analyses, insufficient DNA was available from all replicates of line S3 and it's offspring for 454 pyrosequencing. Therefore, only one DNA replicate from each line was amplified and subjected to pyrosequencing. This did not allow MANOVA to be performed on the data as was the case for the analysis of peak heights of the Bg112 locus. However, it did allow the generation of additional data supporting genetic differences among S3 and its offspring lines using an independent method.

Effect of genetic exchange and segregation on plant growth

Genetic exchange experiment: Parental lines C2, C3 and D1 and crossed lines S1, S3, S5, Sb, Sc2 and Sd were used. Seeds of *Plantago lanceolata* were obtained from Fenaco (Winterthur, Switzerland) and seeds of *Oryza sativa* (Nipponbare) were obtained from Uta Paszkowski (University of Lausanne). The seeds were surface-sterilized in sodium hypochlorite (NaClO) (6% for *P. lanceolata* and 2% for *O. sativa*), washed repeatedly in sterile water and put on filter paper for germination in the greenhouse. Five days after germination for *P. lanceolata*

and eight days for *O. sativa*, seedlings were put on trays with moist vermiculite for another five days. A total of 200 pots were filled with 350 ml of a 1:1 (vol:vol) mixture of loam and sand for *P. lanceolata* and with 450 ml of a 4:6 (vol:vol) mixture of loam and sand for *O. sativa*. The soil mixture was steam autoclaved twice at 120°C two weeks before planting. Seedlings of similar size of either one of the two host species were individually transferred to 100 pots that had been well watered (10 replicate pots for each parental line, each crossed line and non-mycorrhizal treatment (NM) for each plant species). The position of the individual pots was randomized at the beginning of the experiment. Each seedling was inoculated with 0.2 ml of spore suspension that contained 500 spores of one of the parental lines or one of the crossed lines. The NM treatment received 0.2 ml of sterile water. Each pot was watered every two days with 30-50 ml tap water for *P. lanceolata* and with 50-75 ml tap water for *O. sativa*. Each pot was fertilized twice with full-strengh Hoagland solution containing no phosphorous and the K concentration was adjusted by adding KCl. Day length was 16h and the temperatures ranged between 18 and 30°C. After 12 weeks, the shoots of each plant were harvested. The roots were then carefully washed. After recording the fresh weight of the roots, a small root sample was cut and stored in 50% ethanol in order to test for fungal colonization, and the remaining fresh weight of the roots was recorded, allowing us to determine the dry weight of this piece of root. Shoots and roots were separately dried at 80°C for two days and then weighed. All the roots were randomly washed and weighed within the following ten days. All the non-mycorrhizal (NM) plants were free of fungi. Seven plants died during the experiment and in the roots of five plants where inoculum was added, no fungal colonization was observed. These twelve individuals were removed which left data from 188 individuals for the analyses.

Segregation experiment

Crossed lines S3, S4 and Sc2 and the subsequent segregated lines were used. The protocols were the same as before, except for the following. A total of 380 pots (10 replicates per treatment) were filled with 350 ml of a 1:1 (vol:vol) mixture of loam and sand for *P. lanceolata* and with 700 ml of a 4:6 (vol:vol) mixture of loam and sand for *O. sativa*. Each pot was watered every two days with 40 ml tap water for *P. lanceolata* and with 70-100 ml tap water for *O. sativa*. All the non-mycorrhizal (NM) plants were free of fungi. Sixteen plants either died or failed to be colonized during the experiment, which left the data to 364 individuals for the analyses.

Quantification of symbiosis-specific gene transcripts in rice

Spores of crossed line S4 and the subsequent segregated lines were used to inoculate plants of *O. sativa* with same protocol as used in the segregation experiment, except for the following. After 12 weeks, roots were carefully washed and frozen in liquid nitrogen, immediately after the harvest, for subsequent RNA extractions (on mycorrhizal and NM roots). Eleven plants either died or failed to be colonized during the experiment, leaving material from 39 individuals for the analyses. RNA was extracted from 100 mg of ground root tissue using the Rneasy Plant mini kit according to the manufacturer's instructions (Qiagen®). Total RNA was treated with DNase I grad according to manufactured protocol (invitrogen®). To eliminate sugar present in great quantity in plant roots, a LiCl treatment was performed. Absence of genomic DNA was confirmed by PCR on nearly constitutively expressed cyclophilin gene (TIGR identifier, LOC_Os02g02890). First strand cDNA synthesis was carried out using superscript III (Invitrogen®) following the manufacturer's instructions. Real-time PCR was carried out in optical 384-wells plates. A Tekan robot was used for pipetting into the plate to avoid liquid handling errors. Power SYBR® green master mix was used as a label (Applied Biosystems). Fluorescence was quantified with a Prism 7900HT sequence detection system (Applied Biosystems). PCR conditions and mix for each tube was identical to that described by Gutjahr *et al.* (2008). Primer sequences for real-time RT-PCR are shown in Table S5. Transcription values were calculated with SDS2.2.1 software (AB) and qBASE 1.3.5 software (Hellemans *et al.*, 2007). They were normalized to the geometric mean of amplification of the nearly constitutively expressed gene: polyubiquitin (TIGR identifier, LOC_Os06g46770). In addition to four symbiosis specific genes (AM1, AM3, AM14 and Pt11), we also analysed the transcription levels of the housekeeping gene encoding cyclophilin. Seven samples were not used because of the poor quality of the RNA.

Statistical analyses

AFLP: To detect differences among segregated lines, *G*-tests of independence with a Williams's correction (Sokal *et al.*, 1995) were used. Firstly, we performed *G*-test to compare the proportion of polymorphic markers among segregated lines (number of genetically different markers *vs.* number of similar markers compared to crossed lines). We also performed *G*-test to compare the proportion of disappearing and appearing markers among

segregated lines (number of disappearing markers *vs.* number of appearing markers compared to crossed line).

Sequence based markers

To detect segregation, multivariate analyses of variance (MANOVA) were performed on the relative peak heights of the four alleles (amplified with the pair of primers Bg112-up and Bg112-d) and with the main factor AMF lines.

Segregation and genetic exchange on plant growth

Statistical analyses were performed on the dry weight. For each experiment, to test whether the overall growth of NM plants differed from AMF inoculated plants, all the AMF lines were pooled and compared with the NM treatment using one-way ANOVAs. No overall effect of AMF inoculation was found in the segregation experiment. We therefore used Dunnett's tests with an alpha level of 0.05 to compare the NM treatment with each AMF line. To test the effect of genetic exchange on plant growth, the NM treatment was removed because it could bias the results of the statistical analyses. We separated the data obtained for each pairing used to produce the crossed lines as it is not relevant to compare crossed lines of two different pairings. In the same way, we separated data obtained for each crossed line used to produce the segregated lines. We performed crossed two-way ANOVAs on these data with "AMF line" and "plant species" as fixed factors. One-way ANOVAs were then performed for each plant species separately and when ANOVAs were significant, we carried-out a multiple comparison test using Tukey-kramer HSD (honestly significant difference) test with an alpha level of 0.05.

Plant gene transcription: We analysed the transcription level data obtained for the four symbiosis specific genes (AM1, AM3, AM14 and Pt11). Relative plant gene transcription of the NM treatments were around zero (2.02, 8.05, 8.09 and 5.23 for AM1, AM3, AM14 and PT11, respectively). We firstly pooled all the data from AMF lines and we performed a crossed two-way ANOVA with "genes" and "NM/mycorrhizal plants" as fixed factors. We found strong effects of genes and AMF inoculation (Genes: $F_{3,114}$=10.69, $p < 0.001$; NM/mycorrhizal plants: $F_{1,114}$=263.27, $p < 0.001$); highlighting the specificity of the genes to the symbiosis. Because we wanted to detect possible significant differences in symbiosis-specific gene transcripts among AMF lines we removed the NM treatment for subsequent analyses as it would strongly bias the results. One-way ANOVAs were then performed on the

transcript levels for each gene separately among AMF treatments. Where analyses of variance were significant, we carried-out a multiple comparison test using Tukey-kramer HSD test with an alpha level of 0.05. We performed the same analysis on the transcription levels of the cyclophilin gene. Significant differences obtained among AMF lines did not change by adding the NM treatments.

All analyses were performed with the statistical program JMP® version 5.0 (SAS Institute Inc., Cary, NC, USA).

Curriculum Vitae

Professional address
Department of Ecology and Evolution
University of Lausanne
Biophore Building
CH-1015 Lausanne
Switzerland
Tel: +41 (0)21 692 4246

Personal information
Chemin des platane 6
CH-1008 Prilly
Switzerland
Tel: +41 (0)78 645 15 29
Email: colard.alexandre@gmail.com
Born in December 18th 1981

Colard Alexandre

Professional experience

2006 -2011	PhD STUDENT – FNRS fellowship Effect of segregegation and genetic exchange on plant gene expression Ecology and Evolution Department, Lausanne University
Collaboration:	Uta Paszkowski – Molecular and Vegetal Biology Department, Lausanne University Center for Integrative Genomic – Lausanne DNA array Facilities (DAFL), Lausanne University Francis Martin's group – Ecogenomics of Interactions, INRA, Nancy

Education and Training

2006 -	PhD student in Switzerland in "Department of Ecology and Evolution" Under supervising of Ian Sanders, Lausanne, Switzerland **Thesis title**: Effect of segregegation and genetic exchange on plant gene expression
2004-05	Master of Molecular Evolution in Paris XI France (mention Bien) (equivalent to Msc) Training course in laboratory "Population, Génétique, Evolution" CNRS Under supervising of Catherine Montchamp-Moreau – CNRS - Laboratory of Population Genetic and Evolution, Orsay, France **Thesis title**: Etude de gènes candidats pour la distorsion de ségrégation méiotique des chromosomes sexuels chez *Drosophila simulans*
2003-04	Maitrise de Biologie Moléculaire et Physiologie Animale (mention Assez Bien) (MSc 1st year) PARIS VI, France **Thesis title**: Domestication des Éléments transposable chez les mammifères
2003	Training Course: Study of SRF (Serum Response Factor) gene in mouse Institut Cochin Hospital, Paris, France
2002-03	Licence de Biologie Moléculaire et Physiologie Animale (BSc 3rd year), PARIS VI, France
2001-02	DEUG – Science et Vie de la Terre (BSc 1st and 2nd year), Paris VI, France
1999-01	Pharmacology University, Paris V, France
1998-99	Scientific Baccalaureat, High-shool Joliot Curie, France

Teaching

2006-10 Master student supervision (3 month)

Title: Effect of genetic drift in AMF on Plant gene expression

Master practical works of molecular biology (140 hours)

Bachelor 2nd and 3rd year: Bioinformatics practical works (32 hours)

Bachelor 2nd year: Statistical practical works (6 hours)

Bachelor 1st year: Botany and mycology practical works (160 hours)

Grant

2011 Societé Academic Vaudoise – 10 000 CHF

Skills

Project Management	PhD – set up experiment, budget, analysis and results presentation
	Master student training course supervisor – set up experiment, budget, results analysis and presentation
Laboratory techniques	PCR, Reverse transcriptional PCR, Real-time PCR, cloning, sequencing, Northern Blot, DNA & RNA extraction, AFLP, Microsatellite, *in vitro* culture
Computing	MEGA 4.0, BioEdit, Qbase, GeneMapper, Microarray Analysis (R, DAVID, Agrigo), Photoshop, Word, Excel, Powerpoint
Languages	French mother tongue, English fluent

Oral and Poster Communications

2009 **Colard Alexandre**, Angelard Caroline, Ian Sanders
Genetic exchange in AMF alter plant gene expression
Poster presentation
International Conference of Mycorrhyza Belo Horizonte, Brazil

2009 **Colard Alexandre**, Angelard Caroline, Ian Sanders
Effect of genetic exchange on plant gene expression
Oral presentation
Biology09 Bern, Switzerland

2008 **Colard Alexandre**, Angelard Caroline, Ian Sanders
Effect of genetic exchange on plant gene expression
Poster presentation
Biology08, Lausanne, Switzerland

2007 **Colard Alexandre**, Angelard Caroline, Ian Sanders
Effect of genetic exchange alter timing of colonization
Oral presentation
Evolution, Minneapolis, USA

2007 **Colard Alexandre**, Sanders Ian
Do AMF harbour meiosis-specific genes in their genome?
Oral presentation
Biology07, Zurich, Switzerland

2006 **Colard Alexandre**, Sanders Ian
Do AMF harbour meiosis-specific genes in their genome?
Poster presentation
International Conference of Mycorrhyza Granada, Spain

Award

2007 "Ecogenics innovation award 2007" – for the oral presentation at Biology07

Publications

2011 **Colard A.**, Angelard C. and Sanders I.R., Effect of segregation in AMF on plant's transcriptome, *in prep.*

2011 **Colard A.**, Angelard C., Martinez J.J.M., and Sanders I.R., Using natural genetic variation in arbuscular mycorrhizal fungi as an additional approach to identifying genes involved in the mycorrhizal symbiosis, *in prep.*

2011 Roger A, **Colard A.**, Angelard C. and Sanders I.R., Relatedness of arbuscular mycorrhizal fungi drives plant growth via intraspecific fungal competition, *in prep.*

2011 Tisserant E, Kohler A, Seddas P, Balestrini R, Benabdellah K, **Colard A**, Croll D, Gomez SK,Koul R, Ferrol N, Fiorilli V, Formey de Saint Louvent D, Franken Ph, Helber N, Hijri M,Lanfranco L, Lindquist E, Liu Y, Malbreil M, Morin E, Shapiro H, van Tuinen D, Waschke A,Azcón C, Bécard G, Bonfante P, Gianinazzi-Pearson V, Harrison MJ, Paszkowski U, Requena N, Rensing SA, Roux C, Sanders IR, Shachar-Hill Y, Tuskan G, Wincker P, Young P,Lammers P, and Martin F, The transcriptome of the arbuscular mycorrhizal fungus *Glomus intraradices* reveals functional tradeoffs in an obligate symbiont, *submitted*

2011 **Colard A.**, Angelard C., Sanders I.R., Effect of genetic exchange on mycorrhizal specific genes expression. *Applied and Environmental Microbiology, in press*

2010 Angelard C., **Colard A**., Niculita-Hirzel H., Croll D., Sanders I.R., Segregation in a mycorrhizal fungus alters rice growth and symbiosis-specific gene transcription. *Current Biology*

2009 Corradi N., Ruffner B., Croll D., **Colard A**., Horák A., Sanders I.R. High-level molecular diversity of copper-zinc superoxide dismutase genes among and within species of arbuscular Mycorrhizal fungi. 2009 *Applied and Environmental Microbiology* 75(7), 1970-1978

2007 Corradi N., Croll D., **Colard A**., Kuhn G., Ehinger M., Sanders I. R. Gene copy number polymorphisms in an arbuscular mycorrhizal fungal population. *Applied and Environmental Microbiology* 73(1), 366-9

2006 Montchamp-Moreau C., Ogereau D., Chaminade N., **Colard A**., Aulard S. Organization of the sex-ratio meiotic drive region in Drosophila simulans. *Genetics* 174(3), 1365-1371