

Università degli Studi di Torino

Tesi di Dottorato di Ricerca in Scienze Biologiche
e Biotecnologie Applicate

PhD Thesis in Biological Sciences
and Applied Biotechnologies



**Analysis of RNA interference and small RNAome in
arbuscular mycorrhizal fungi**

Alessandro Silvestri

Tutor: Prof. Luisa Lanfranco

XXXII Cycle: 2016 – 2019

PREFACE	3
1. General introduction	7
1.1. The discovery of the RNA interference	8
1.2. The core of RNAi machinery: DCL, AGO and RdRp	8
1.3. Different classes of sRNAs are involved in RNAi	10
1.4. A focus on fungal RNAi	14
1.5. Fungal RNAi in antiviral defense	18
1.6. The role of sRNAs and RNAi in inter-kingdom interactions	19
1.7. The Arbuscular Mycorrhizal Symbiosis	22
1.7.1. The fungal partner	23
1.7.2. The colonization process	24
1.7.3. Flux and role of nutrients (phosphorus and carbon) in AMS	26
1.7.4. An emerging role for fungal effectors, peptides and small RNA in AMS	29
1.8. Aims of the PhD thesis	32
1.9. References	33
2. In silico analysis of fungal small RNA accumulation reveals putative plant mRNA targets in the symbiosis between an arbuscular mycorrhizal fungus and its host plant	47
3. Rhizophagus irregularis sRNA Rir-2216 is likely involved in cross-kingdom RNAi by silencing host plant gene	67
3.1. Introduction	68
3.2. Results and Discussion	69
3.3. Conclusions and Perspectives	76
3.4. Materials and Methods	77
3.5. References	81
4. The virome of the arbuscular mycorrhizal fungus Gigaspora margarita reveals the first report of DNA fragments corresponding to replicating non-retroviral RNA viruses in fungi	85

5. Different genetic sources contribute to the small RNA population in the arbuscular mycorrhizal fungus <i>Gigaspora margarita</i>	101
5.1. Abstract	102
5.2. Introduction	102
5.3. Methods	104
5.4. Results and Discussion	107
5.5. Conclusion	125
5.6. References	126
6. General remarks and perspectives	133
6.1. General remarks and perspectives	134
6.2. References	137

PREFACE

My thesis work has been focused on the role of small RNAs (sRNAs) and RNA interference (RNAi) in Arbuscular Mycorrhizal Symbiosis (AMS), an ancient and widespread mutualistic symbiotic association formed between the roots of the majority of land plants and a group of fungi, known as Arbuscular Mycorrhizal Fungi (AMF). AMS is based on nutrient exchange between the symbiotic partners: plants provide up to 20% of their own fixed carbon to AMF in exchange of mineral nutrients collected by fungal extra-radical hyphae which are very efficient in the exploration and exploitation of soil resources. The establishment of AMS results in an increased fitness for the host plants and, at the same time, allows the AMF to complete the life cycle.

The symbiotic nature of AMS, that arises from a long history of co-evolution, requires an extremely complex communication between the symbionts that makes use of a number of signaling molecules, including hormones, cell wall-derived compounds and secreted proteins. However, despite enormous advances achieved in this direction in recent years, the complete picture of this molecular communication is far to be fully understood. Interestingly, sRNAs are currently emerging as key components in gene expression regulation at the basis of several biological processes in eukaryotes as well as important regulators in different inter-species, and even inter-kingdom, interactions. Small RNAs are short non-coding RNA molecules that, through the RNAi pathway, are primarily known to suppress the expression of endogenous genes that are recognized through sequence complementarity. It has been recently demonstrated that sRNAs can also be transferred from a “donor” to a “receiver” organism where they can regulate host gene expression in a process named “cross-kingdom RNAi”. Whether this phenomenon occurs in AMS is the main issue that I investigated in my PhD research project, in particular focusing on the hypothetical movement of sRNAs from AMF to the host plants. Considering that at the beginning of this project almost nothing was known about the “sRNA world” in AMF, I have started with the characterization of the fungal RNAi molecular components and sRNA populations in AMF. In the thesis, I will provide some preliminary evidence of their biological functions.

This thesis is divided in six chapters. In **Chapter 1**, I introduced the main topics instrumental to follow the next sections. At first, I described the general

aspects of RNAi pathway and the classification of sRNAs, using plants as model systems. I then made a focus on what is known on RNAi pathways in the fungal kingdom. Finally, I presented the most recent advances on cross-kingdom RNAi as an inter-species communication mechanism and introduced the biological system on which I worked on, the AMS.

In **Chapter 2**, I provided the first description of the RNAi machinery and the sRNA population for an AMF, the model species *Rhizophagus irregularis*. This fungus is equipped with a peculiar RNAi machinery characterized by single Dicer-like (DCL) and an unusual high number of Argonaute-like (AGO-like) and RNA-dependent RNA polymerase (RdRp) genes, which was later found to be a common feature of AMF. Furthermore, the analysis of sRNA population highlighted the existence of more than one sRNA-generating pathways that are likely differentially modulated between extra-radical and intra-radical mycelium. I also reported *in silico* evidence, by means of a degradome-seq analysis, of host plant genes potentially silenced by *R. irregularis* sRNAs.

Chapter 3 reports preliminary experimental validations of some of the predictions identified in Chapter 2. In particular, I demonstrated that the *R. irregularis* sRNA *Rir2216* is able, in an heterologous system, to silence a plant host WRKY transcription factor, which was also found to be down-regulated in specific root cortical cells upon mycorrhization. This is the first experimental evidence suggesting the occurrence of cross-kingdom RNAi in AMS.

Chapter 4 describes the contribution to the characterization of the virome of the AMF *Gigaspora margarita*, a species hosting also a population of endobacteria. This AMF is infected by four mitoviruses, one Ourmia-like narnavirus and one Giardia-like virus. In this work we reported, for the first time outside the animal kingdom, the occurrence of DNA fragments corresponding to non-retroviral replicating RNA viruses. Interestingly, a similar mechanism has been characterized in insects as an anti-viral defense strategy.

In **Chapter 5**, I investigated the RNAi machinery and the sRNAome of *G. margarita*. This species, that is phylogenetically distantly related to *R. irregularis*, shows a core of RNAi-related genes similar to that of *R. irregularis*, but a substantially different population of nuclear DNA mapping sRNAs. I also characterized all the different genetic sources giving rise to sRNAs considering the nuclear, mitochondrial, endobacterial and viral genomes. In this analysis I highlighted specific elements that suggest an active role for *G. margarita* RNAi in anti-viral defense. I also reported evidence of a potential involvement of

viral-derived *G. margarita* sRNAs in the silencing, through cross-kingdom RNAi, of host plant genes.

In **Chapter 6** I ended the thesis with some general conclusions and perspectives.

CHAPTER 1

1. General introduction

1.1. The discovery of the RNA interference

RNA interference (RNAi) is a biological process, almost universally presents in eukaryotes, involved in regulation of gene expression at transcription and/or post-transcription levels. The first reported observations of RNAi date back to the early 1990s when two independent research groups, that were working on the over-expression of genes involved in pigment production in a plant (Napoli et al., 1990) and a fungus (Romano and Macino, 1992), obtained unexpected results. In particular, they observed that a number of transformants, instead of being characterized by a more intense pigmentation, showed albino or variegated phenotypes. The peculiar feature was that those phenotypes were often transient and not always transmitted to the progeny; they called this phenomenon as “co-suppression” and “quelling”, respectively. The molecular basis of this mechanism remained unclear until 1998 when Fire and colleagues understood that the nature of the silencing trigger was double-stranded RNA (dsRNA, Fire et al., 1998). Their work brilliantly highlighted that, in nematodes, dsRNAs activate the sequence-specific silencing of endogenous genes, producing a much stronger effect compared to sense or antisense sequence-specific single-stranded RNAs (Fire et al., 1998). During the following years, the molecular details of RNAi were elucidated giving rise to a number of biotechnological applications, from medicine to agriculture. For the impact of this discovery, in 2006 Fire and Mello were awarded the Nobel Prize in Physiology or Medicine.

1.2. The core of RNAi machinery: DCL, AGO and RdRp

In a simplified model, the general RNAi pathway can be described as a process involving three classes of proteins: Dicer (or Dicer-like; DCL), Argonaute (AGO) and RNA-dependent RNA polymerase (RdRp; Ipsaro and Joshua-Tor, 2015).

DCL are enzymes that process dsRNAs or single-stranded hairpin RNAs into small RNA (sRNA) duplexes, typically 21-25 nucleotide(nt)-long (Bernstein et al., 2001). The cleavage activity of DCL is provided by a pair of Ribonuclease III (RNase III) domains which act, together with a PAZ domain (primarily responsible for the recognition of the dsRNAs to be cleaved), as a molecular

ruler to produce sRNA duplexes of the appropriate size, characterized by 2-nt overhangs at the 3'-ends and by a monophosphate group at the 5'-ends (Zhang et al., 2004; MacRae and Doudna, 2007). The sRNA duplexes are then loaded onto AGO, the silencing effector proteins, which retain only one sRNA strand to become the minimal constituents of the (1) RNA-induced silencing complex (RISC; Cenik and Zamore, 2011), involved in post-transcriptional gene silencing (PTGS), or the (2) RNA-induced transcriptional silencing complex (RITS; Moazed, 2009), a nuclear form of RISC involved in transcriptional gene silencing (TGS). The sRNAs let the RISC or RITS to recognize, by sequence complementarity, their target nucleic acid molecules (RNA or DNA, respectively) that can therefore be silenced in different ways: (1) through transcription suppression, mRNA cleavage or destabilization, or translational repression (Verdel et al., 2004; Ipsaro and Joshua-Tor, 2015); or (2) through histone and/or RNA-directed DNA methylation (Moazed, 2009).

AGO have a bilobal structure: one lobe is composed by a PAZ (the same found in DCL) and a N-terminal domains, the other by a MID and a PIWI domains (Kuhn and Joshua-Tor, 2013). The PAZ/N-terminal lobe recognizes the DCL-derived sRNA duplexes, thanks to their typical 2-nt 3'-overhangs, and binds only one strand of the sRNA duplex (the sRNA guide) at its 3'-end, discarding the other one. The second lobe, MID, is responsible for securing the sRNA guide to the protein by the aspecific binding of its 5' monophosphate, while the PIWI domain, which possesses an RNase-H fold, provides the RISC complex with the endonuclease catalytic effect directed against the target transcripts. Three paralogous classes of AGO are present in eukaryotes: the widespread AGO-like, the Piwi-like (only found in animals) and the *Caenorhabditis elegans*-specific group 3 AGO (Hutvagner and Simard, 2008; Wilson and Doudna, 2013). AGO are considered as the minimal and necessary determinant of RNAi; indeed, unlike DCL and RdRp, no AGO-independent RNAi pathways have been described so far (Ipsaro and Joshua-Tor, 2015).

The third class of proteins, the RdRP, is responsible for the “cascade effect”: a fundamental property of RNAi through which a weak initial silencing signal can be drastically amplified in a positive feedback process (Baulcombe, 2015). In this way, few silencing RNA molecules can produce a massive silencing response, resulting sometimes in a huge reprogramming of the cell transcriptome. RdRp are commonly present in fungi and plants but absent,

with the exception of nematodes, in the majority of animals. The RdRp-mediated silencing amplification can be accomplished in different ways: (1) RdRp can interact with AGO-cleaved mRNAs to synthesize their antisense fragments (in a sRNA primer-independent or -dependent manner), producing dsRNAs that are then recruited in the DCL circuit; or (2) they can directly produce sRNAs using an mRNA as a template. In plants RdRp are fundamental for defense against viral infections (Ghildiyal and Zamore, 2009). In the simplified model proposed above, several accessory molecular components have been (intentionally) omitted. At molecular level, all these components cooperate to produce a complex and variegated system that can be differentiated, following a reductionist perspective, in a number of RNAi sub-pathways (Ghildiyal and Zamore, 2009).

1.3. Different classes of sRNAs are involved in RNAi

Small RNAs are the informational components of RNAi, directing its activity in a sequence-specific manner. Different classification systems have been proposed for sRNAs; anyway, because of the intrinsic variegated nature of RNAi, a classification system valid for the whole eukaryotic kingdom is missing. Here it is useful to briefly introduce the classification system proposed by Axtell (2013) for plant endogenous sRNAs. This model, that is partially valid also for non-plant sRNAs, contemplates a hierarchical structure on three levels (Fig. 1). The first one distinguishes the sRNAs on the basis of their RNA precursors in (1) hairpin sRNAs (hpRNAs), which include all sRNAs deriving from single-stranded RNA molecules with an intra-molecular self-complementary structure and in (2) small-interfering RNAs (siRNAs), which include all those sRNAs deriving from double-stranded RNA precursors. The second hierarchical level differentiates the (1) hpRNAs in (1A) microRNAs (miRNAs) and (1B) every other hpRNAs that do not qualify as miRNA (other hpRNAs) and the (2) siRNAs in (2A) heterochromatic siRNAs (hetsiRNAs), (2B) secondary siRNAs and (2C) natural antisense transcript siRNAs (NAT-siRNAs):

-1A) miRNAs are defined by their specific and well-characterized biogenesis (Axtell and Meyers, 2018). Genes encoding for miRNAs are initially transcribed by RNA polymerase II in poly-adenylated hpRNA precursors, known as primary miRNAs (pri-miRNAs), which are then consecutively

processed by DCL to originate precursor miRNAs (pre-miRNAs) and mature miRNA duplexes, respectively. A miRNA duplex is constituted by an active strand (the mature miRNA) and a complementary strand (known as miRNA*). Plant miRNAs are generally 20- to 22-nt long and are involved in post-transcriptional gene silencing (PTGS) through the cleavage or the translational repression of the RNA target sequences (Cuperus et al., 2011). Remarkably, most of miRNA families, and their targets, are phylogenetically conserved in the plant kingdom (Jones-Rhoades, 2012).

-1B) other hpRNAs represent all the sRNAs originating from single-stranded hairpin, generally longer than pri-miRNAs, that do not meet the criteria for being annotated as miRNAs (Axtell and Meyers, 2018). The knowledge of this class of sRNAs is still extremely limited.

-2A) hetsiRNAs are generally 23-24-nt long and derive from single-stranded short RNA precursors (~30 to 50 nt) transcribed by the RNA polymerase IV from intergenic, repetitive regions and/or transposable elements (Matzke and Mosher, 2014). Their precursors are converted to dsRNAs by RdRp and cleaved by DCL. HetsiRNAs are the most abundant sRNAs in plants and are involved in transcriptional gene silencing (TGS) through histone and/or RNA-directed DNA methylation.

-2B) secondary siRNAs derive from RNA precursors initially transcribed by RNA polymerase II from non-coding regions, transposable elements or protein-encoding genes that, after a siRNA- or miRNA-directed AGO-catalyzed cleavage event, are converted in dsRNAs by specific RdRp. Sometimes the dsRNA production is sRNA-independent, but dependent on spliceosome (Christie et al., 2011) or other RNA processes, such as RNA decay (Marí-Ordóñez et al., 2013). The dsRNAs are then processed by DCL to originate 21- or 22-nt long secondary siRNAs, often in a “phased” manner (phasiRNAs) from consistent dsRNA ends (Fei et al., 2013). Secondary siRNAs may have *cis* or *trans* PTGS targets whose cleavage can trigger new secondary siRNA production; thus, they are generally included in complex silencing cascades able to regulate, in a coordinated way, large gene families, such as the plant disease-resistance NB-LRR (Fei et al., 2013).

-2C) NAT-siRNAs are the only class of plant siRNAs that are independent from RdRp. Indeed, they are produced by the DCL cleavage of dsRNAs that originate from the hybridization of transcripts encoded by opposite strands of the same genomic region (*cis*-NAT-siRNA; Henz et al., 2007). Theoretically, NAT-siRNAs can also derive from the hybridization of transcripts that do not possess a genomic overlap (*trans*-NAT-siRNAs), although this case has never been reported so far (Axtell, 2013).

A third hierarchical level for the classification of plant sRNA has also been proposed (Fig. 1): it includes long miRNA but, it is likely valid only for some plant species (Axtell, 2013).

A further important class of sRNA that worth to be mentioned is the Piwi-associated RNAs (piRNAs): these sRNAs are present only in animals and are processed in a DCL- and RdRp- independent manner (Ipsaro and Joshua-Tor, 2015). The piRNA precursors are transcribed from peculiar loci, called piRNA clusters, are then processed in the cytosol by various nucleases (such as Zucchini) into 26- to 31-nt long sequences and are finally loaded onto the animal-specific Piwi-clade AGO (Juliano et al., 2011). A hallmark of piRNA pathway is the ping-pong amplification cycle in which, through Piwi-clade AGO- or Aubergine- and AGO3-dependent manner, secondary piRNAs are produced.

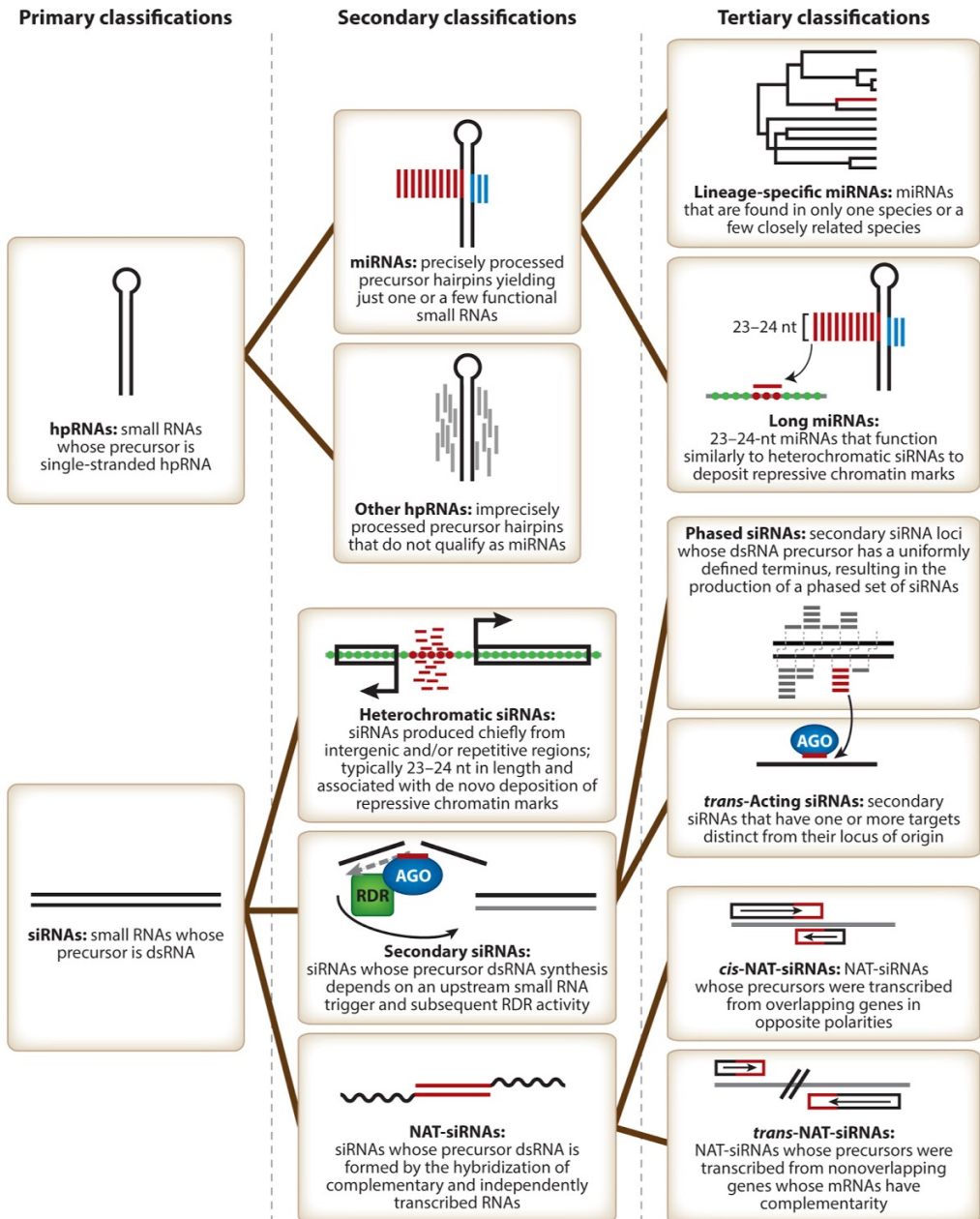


Figure 1. Hierarchical classification system for endogenous plant small RNAs (from Axtell, 2013). Abbreviations: dsRNA, double-stranded RNA; hpRNA, hairpin RNA; miRNA, microRNA; NAT-siRNA, natural antisense transcript small interfering RNA; siRNA, small interfering RNA.

miRNAs, siRNAs and piRNAs are regulated *in vivo* by specific modifications (Borges and Martienssen, 2015). Among them, two of the best characterized are the uridylation and the 2'-O-methylation of the 3'-ends of sRNAs, which have two opposite molecular meanings (Kim et al., 2010). In plants, the 2'-O-methylation of sRNAs, which is carried out by the enzyme HEN1, is associated with sequence stability: indeed, the methylation prevents sRNAs from being uridylated by HESO1 and, consequently, from being degraded by SDN1 (Borges and Martienssen, 2015). For this reason, functional sRNAs are generally methylated during their biogenesis. HESO1 is associated with the RISC complex where it promotes (and this seems to be its primary function) the uridylation, and thus degradation, of the 3'-ends of the 5' fragment of the AGO-cleaved mRNAs (Ren et al., 2014). In this context, the 2'-O-methylation is necessary to avoid the unintentional uridylation by HESO1 of the AGO-loaded sRNA guide that would lead, after its degradation, to a loss of silencing information.

1.4. A focus on fungal RNAi

RNAi is a widespread mechanism within the fungal kingdom; it is known that almost all fungal lineages possess the key components of RNAi machinery (reviewed in Nicolás and Garre, 2017; Torres-Martínez and Ruiz-Vázquez, 2017). However, the molecular details of its functioning have been deeply characterized only in few model species and, in particular, concerning its role in the protection of genome integrity. The emerging picture suggests that, in general, fungal RNAi seems to be, more than in plants, a highly variegated mechanism that even includes a number of species-specific (sub-)pathways. Some of these pathways will be discussed below.

The first characterized fungal RNAi pathway was the “quelling” phenomenon in *Neurospora crassa* (Romano and Macino, 1992). Quelling is a PTGS pathway involved in genome defense, through the suppression of transposon activity during the vegetative growth; it relies on a DNA/RNA-dependent RNA polymerase (D/RdRp; named QDE-1), on two functionally redundant DCL (DCL-1 and DCL-2) and an AGO (QDE-2). QDE-1 synthesizes single-stranded aberrant RNAs from tandem repeats - typical of transposable elements - and

turns them into dsRNAs. The latter are then processed in 25-nt long siRNAs by DCL-1/2 and loaded onto QDE-2 for the PTGS of RNA target sequences.

In *N. crassa*, also the qiRNA (QDE-2-interacting small RNA) pathway participates to the genome defense during vegetative growth (Lee et al., 2009). This system is mechanistically similar to quelling but it distinguishes from the latter mainly because it gives rise to specific sRNAs (qiRNAs) only from repetitive ribosomal DNA clusters. qiRNAs are produced as a response to double-strand DNA breaks and likely act as inhibitors of protein translation and so as a defense strategy to protect the cell until the DNA damage is repaired.

Quelling and qiRNA pathways are both dependent, for their activation, on homologous recombination (HR). Indeed, the HR is the molecular mechanism exploited by the cell allowing it to discriminate the repetitive regions from the rest of the genome, since the latter is much less prone to undergone HR (Dang et al., 2014). Contrary to the multiple transgenic loci often originated by genetic transformation, the ribosomal DNA loci, despite their repetitive nature, are generally protected from HR in the absence of DNA damage to prevent the loss of rDNA copies (Dang et al., 2014). In this context, a double-strand DNA break is the necessary determinant to trigger the qiRNA pathway.

A further characterized RNAi pathway in *N. crassa* participates to the protection of genome integrity during sexual reproduction (Shiu et al., 2001). This process, known as meiotic silencing by unpaired DNA (MSUD), recognizes and silences, during meiosis, the unpaired sequences present only in one parental chromosome (likely mobile elements). In this pathway, a number of proteins initially cooperate to produce aRNAs from the unpaired sequences which are then subsequently converted into dsRNAs by the RdRp SAD-1 and into MSUD-associated siRNAs (masiRNAs; 25-nt long and enriched in uracil at their 5'end) by DCL-1 (Shiu et al., 2006). The AGO protein SMS-2 is finally involved in the silencing of target sequences.

Besides *N. crassa*, the genome protection role of fungal RNAi - in the fight against transposons, retro-transposons, viruses and transgenes - has been characterized also in *Mucor circinelloides*, *Trichoderma atroviridaem*, *Fusarium graminearum*, *Aspergillus nidulans*, *Cryptococcus neoformans* and *Magnaporthe oryzae* during vegetative growth (quelling-like pathways) and in *C. neoformans* during sexual reproduction (Torres-Martínez and Ruiz-Vázquez,

2017 and references therein). Remarkably, a specificity of the quelling-like mechanism of *M. circinelloides* is that, contrary to all fungal systems listed above, it requires two non-redundant RdRp, one of which is involved in the amplification of silencing signals (i.e. the production of secondary siRNAs; Calo et al., 2012). Furthermore, in *Schizosaccharomyces pombe* it has been demonstrated that RNAi is involved in the sequence-specific heterochromatin formation by means of histone modifications (Verdel et al., 2004): this is the only example so far described in the whole fungal kingdom.

As mentioned above, the regulatory role of RNAi in endogenous gene expression has been less widely characterized in fungi compared to other organisms, also because fungal RNAi knock-out mutants often lack evident physiological or developmental phenotypes (Nicolás and Garre, 2017; Torres-Martínez and Ruiz-Vázquez, 2017). Most of the knowledge arises from studies conducted on *N. crassa* and *M. circinelloides*.

A first group of fungal endogenous regulatory sRNAs, the miRNA-like, has been initially identified in *N. crassa* (Lee et al., 2010). These miRNA-like share several features with plant miRNAs: they originate from precise excision of sRNAs from hairpin precursors that are transcribed from non-coding regions. In *N. crassa*, miRNA-like are generally 25-nt long and enriched in uracil at their 3'ends; however, their target sequences are still unknown (Lee et al. 2010). Evidence of miRNA-like has also been collected in other Ascomycota and Basidiomycota species, but never in basal fungi, such as *M. circinelloides* (Villalobos-Escobedo et al., 2016).

A second group of endogenous sRNAs are the exogen-derived siRNAs (ex-siRNAs) of *M. circinelloides* (reviewed in Torres-Martínez and Ruiz-Vázquez, 2016). The ex-siRNAs, that can be divided in 4 classes, originate from mature mRNAs in a DCL-dependent way. Class 1 and 2 ex-siRNAs are generally 23-24 nt-long sequences, are enriched in uracil at their 5'-ends and do not show a strand bias (they are both sense and anti-sense to the parental mRNA). These sRNAs are involved in the suppression of gene expression of the transcripts from which they originate. Class 3 and 4 ex-siRNAs are instead not characterized by specific length in nucleotide, are mostly sense to the originating exons and are not enriched in uracil at their 5'-ends; despite they are DCL-dependent, all these features suggest that they play a role in a non-canonical RNAi pathway.

M. circinelloides is also characterized by another group of gene regulator sRNAs, named rdRNAs, that are produced in a DCL-independent but RdRp-dependent way (reviewed in Torres-Martínez and Ruiz-Vázquez, 2016). These sRNAs share some features with class 3 ex-siRNAs: indeed they are almost all sense to mRNAs, show a random size distribution in terms of nucleotides and are not enriched in uracil at their 5'-ends but at the penultimate position of their 3'-ends. For these reasons, rdRNAs seem to derive from non-random degradation of mRNAs and likely share a partially similar biogenesis pathway with class 3 ex-siRNAs. Interestingly, a new class of RNase III protein (R3B2) displaying unusual domain organization is involved in the biogenesis of rdRNAs. R3B2 have been only found in Mucorales and, for their similarity with RNase III domains of bacteria belonging to Burkholderiales, they are probably a product of horizontal gene transfer event occurred between bacteria and a Mucorales ancestor (Trieu et al., 2015).

Other classes of fungal sRNAs include the NAT-siRNAs (Donaldson and Saville, 2012) and the Dicer-independent esRNAs (disiRNAs; Dang et al., 2013), the latter involved in *N. crassa* in transcription-dependent DNA methylation.

It is worth noting that some fungal species, such as the model yeast *Saccharomyces cerevisiae*, lack key components of RNAi (reviewed in Nicolás et al., 2013 and Nicolás and Garre, 2017). Notably, in *S. cerevisiae* the loss of RNAi genes has been associated with the consistent infection of vertically transmitted dsRNA viruses, known as “killer viruses” (Drinnenberg et al., 2011). These viral entities, which are incompatible with an active RNAi able to suppress their expression, encode for toxic proteins that kill nearby virus-free cells while providing immunity to the virus-infected ones. Under these specific circumstances, the lack of functional RNAi genes can result in a net selective advantage for the virus-hosting individuals. In this context, it is plausible that the original yeast ancestors who lost the RNAi competency were positively selected during evolution, as the killer phenotype more than offsets the disadvantage of lacking RNAi. This phenomenon has been also reported in some basidiomycetes, such as *Ustilago maydis* (Drinnenberg et al. 2011). Moreover, the lack of RNAi in some species is sometimes related to other benefits (Nicolás and Garre, 2017). This is the case of the animal pathogen

Cryptococcus deuterogattii which, because of its strictly parasitic lifestyle, needs to continuously evolve in order to escape the host defense; the lack of RNAi increases the genome mutation rate (due to retro-transposons activity) giving rise to higher genetic variability that can be exploited by the pathogen as a counter-defense strategy (Feretzaki et al., 2016).

1.5. Fungal RNAi in antiviral defense

One of the first functions assigned to RNAi was the defense against viral infections (Nicolás et al., 2013). Fungi from all phyla are indeed infected by viruses (Herrero et al., 2009). Fungal viruses (mycoviruses) are, in most cases, RNA entities encoding for at least one RdRp that is necessary for their replication (Ghabrial et al., 2015). Mycoviral infections are generally asymptomatic (Roossinck, 2011) even if sometimes they can give rise to peculiar phenotypes, such as the case of *Cryphonectria parasitica*, the phytopathogenic agent of the chestnut blight, which, if is infected by viruses of the genus *Hypovirus*, shows a reduced virulence against its host plant (Milgroom and Cortesi, 2004). It is common for fungal endosymbionts of plants to be infected by mycoviruses, the latter can sometimes play mutualistic roles (Bao and Roossinck, 2013). An emblematic example is the tripartite mutualistic interaction between a plant, an endophytic fungus and its corresponding mycovirus, where the mycovirus is necessary to confer heat tolerance to the whole biological system (Márquez et al., 2007).

The replication of almost all mycoviruses is characterized by a dsRNA intermediate; thus, during this phase, host DCL can process the sequences into viral sRNAs able to target the viral genome. A proof that fungal RNAi provides antiviral defense arises from the amount of new mycoviral genomes that have been discovered in the last years by exploiting small RNA-seq data (Nerva et al., 2016). Nevertheless, the sRNA-seq approach is sometimes not sufficient to reconstruct *in silico* the whole genome sequence of mycoviruses, suggesting that some species can escape the host RNAi pathway (Nerva et al., 2018). Anyway, direct evidence of antiviral activity for fungal RNAi has been collected. This is the case of *C. parasitica* infected by *Cryphonectria hypovirus 1* (Segers et al., 2007), *A. nidulans* infected by different viruses (Hammond et al., 2008) and *Colletotrichum higginsianum* infected by *Colletotrichum higginsianum nonsegmented dsRNA virus 1* (Campo et al.,

2016). In analogy to viruses infecting non-fungal hosts, some mycoviruses have developed strategies to suppress the fungal anti-viral RNAi pathway (Segers et al., 2006; Hammond et al., 2008; Yaegashi et al., 2013); these are a further indication that fungal RNAi serves as an antiviral defense system.

1.6. The role of sRNAs and RNAi in inter-kingdom interactions

In recent years it has been discovered that sRNAs play a role, beside endogenous gene regulation and genome protection, also in interspecies communication, in a process known as cross-kingdom RNAi (reviewed in (Chaloner et al., 2016; Huang et al., 2019). Cross-kingdom RNAi was firstly observed in the pathogenic interaction between *Arabidopsis thaliana* and the necrotrophic fungus *Botrytis cinerea* (Weiberg et al., 2013). In this work, the authors brilliantly demonstrated that, during infection, some 21-22 nt-long transposon-derived *B. cinerea* sRNAs are transferred into the host cells (Fig. 2). Once in the plant cells, the fungal sRNAs are loaded onto plant AGO1 which silences, in a PTGS manner, specific defense-related genes, such as mitogen activated protein kinases (MAPKs), cell-wall-associated kinases, and genes involved in the accumulation of reactive oxygen species (ROS); in other words, the pathogen “hijacks” the host RNAi system at its own advantage.

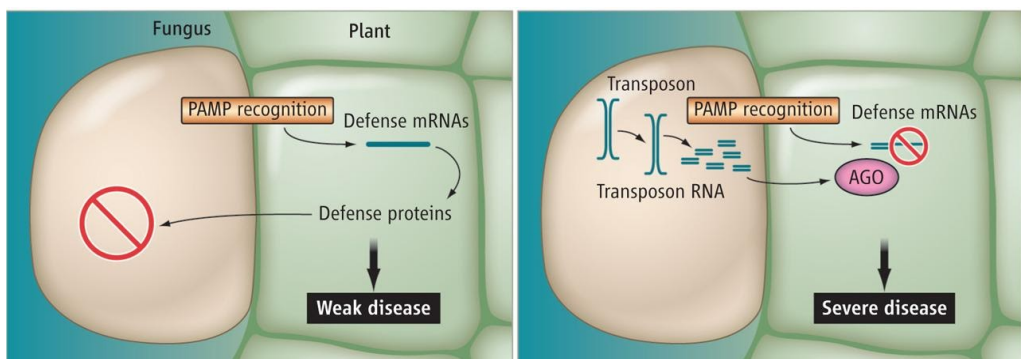


Figure 2. Fungal sRNAs can act as pathogen effectors (from Baulcombe 2013). Plants recognize the presence of fungal pathogens and activate the defense responses (left). Some fungal species can secrete sRNAs into the plant host cells, such as the case of transposon-derived *B. cinerea* sRNAs (Weiberg et al., 2013; Wang et al., 2017), silencing the host defense-related genes (right). These fungal sRNAs act as pathogen effectors by binding the plant AGO proteins.

Since this first discovery, cross-kingdom RNAi has been reported in several other pathogenic/parasitic interactions, in which sRNAs can behave both as “attack” and as “defense” molecules (i.e. can be transferred from the pathogen/parasite to the host and/or *vice versa*; Fig. 3). For examples the fungal pathogens *Verticillium dahliae* (Wang et al., 2016) and *Puccinia striiformis* f. sp. *tritici* (Wang et al., 2017a) and the parasitic plant *Cuscuta campestris* (Shahid et al., 2018) exploit cross-kingdom RNAi to attack their host plants while, at the same time, plants can defend themselves from pathogens by delivering sRNAs, as it has been observed in plant interactions involving *V. dahliae* (Zhang et al., 2016), *B. cinerea* (Cai et al., 2018), *Fusarium graminearum* (Zhang et al., 2016; Jiao and Peng, 2018) and the oomycete *Phytophthora capsici* (Hou et al., 2019).

The occurrence of cross-kingdom RNAi-like mechanisms is not limited to plants. For example, the movement of sRNAs able to interfere with host gene expression has been also reported in the interactions between gastrointestinal nematodes (Buck et al., 2014; Chow et al., 2019) and mammals and between endobacteria (genus *Wolbachia*; Mayoral et al., 2014) or the fungal pathogen *Beauveria bassiana* (Cui et al., 2019) and insects. Furthermore, it has been shown that mobile sRNAs can also regulate gene expression in organism not provided with a functional RNAi machinery, such as the case of the movement of sRNAs from human cells to the malaria agent *Plasmodium falciparum* (LaMonte et al., 2012) (Lamonte et al., 2012) or gut bacteria (Liu et al., 2016). Interestingly, gut microbiome is also influenced by dietary-derived plant sRNAs (Teng et al., 2018). Recently, for the first time it has been demonstrated the involvement of cross-kingdom RNAi also in mutualistic interactions (Ren et al., 2019), as bacteria of genus *Rhizobia* secrete transfer RNA-derived sRNAs into cells of the host legume plant; these sRNAs, by binding plant AGO1, regulate the expression of nodulation-related host genes.

In 2018, Cai et al. (2018) clarified the mechanism through which plants send sRNAs into interacting microbes. In this work, it was shown that, at the *Arabidopsis*-*B. cinerea* contact interface, specific plant exosomes containing sRNAs are secreted into the apoplastic space and are then taken up by fungal cells. The plant knock-out mutants impaired in the production of exosomes were no longer able to defend themselves from the pathogen by means of sRNAs. Similarly, this exosome-mediated transfer mechanism has been reported also in a plant-oomycete interaction (Hou et al., 2019) and in animal

systems (Buck et al., 2014; Chow et al., 2019). In particular, Chow et al. (2019) discovered that the nematode *Heligmosomoides baker* uses extracellular vesicles (EVs) to deliver, into the host cells, sRNAs together with AGO proteins; in this way, it is possible that few transferred sRNA molecules can have a very efficient effect in the recipient organism since it is known that AGO-sRNA complexes can be stable for more than 3 weeks and, during this time, can maintain their silencing competency (Smibert et al., 2013).

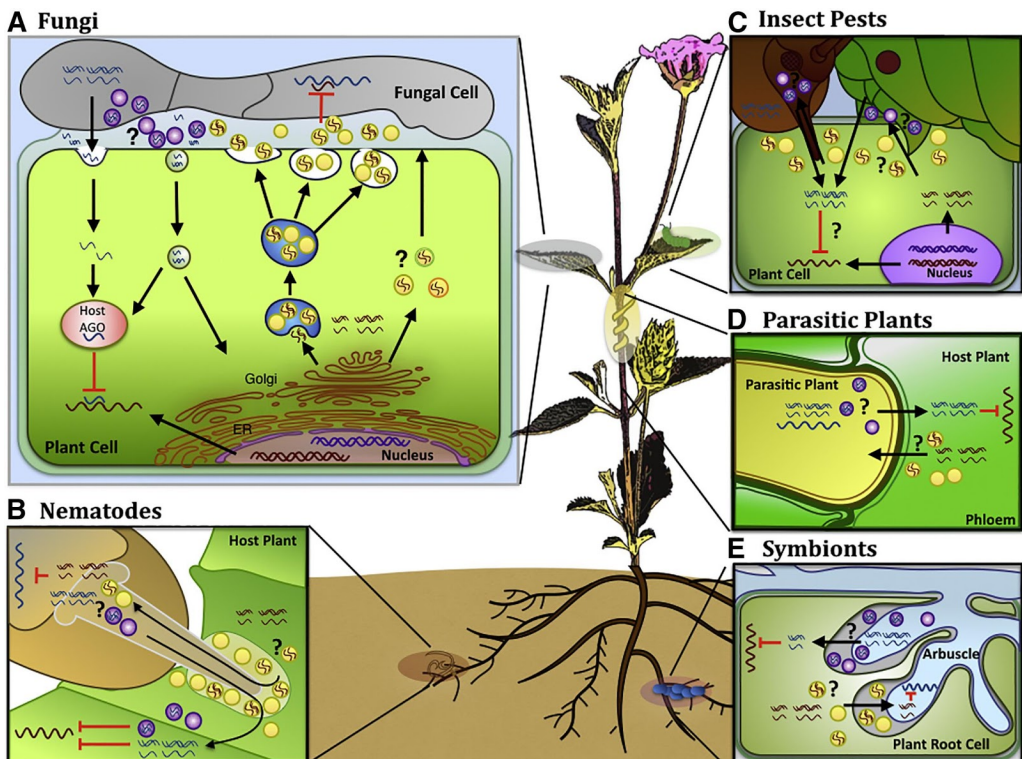


Figure 3. Overview of cross-kingdom RNAi-like processes involving plants (from Huang et al., 2019). The movement of silencing sRNAs has been characterized as a mechanism occurring in plant-fungal pathogen interactions in both directions; the plant-derived sRNAs are transferred into the pathogens through secreted extracellular vesicles (EVs; **A**). Parasitic plants also deliver active sRNAs into host plants to promote colonization (**D**). Cross-kingdom RNAi has also been indirectly demonstrated, through the application of HIGS and VIGS, in the interactions between plants and nematodes (**B**), insect pests (**C**) and Arbuscular Mycorrhizal Fungi (AMF; **E**). It has been speculated that also in these other interacting systems, the sRNAs are exchanged between organisms thanks to EV-mediated processes.

Cross-kingdom RNAi can also be artificially induced through the expression of RNAi constructs directed against interacting organisms, techniques known as host-induced (HIGS) or viral-induced (VIGS) gene silencing (reviewed in Baulcombe, 2015; Fig. 3). Both of these approaches are successful tools exploitable to protect plants from pathogens/parasites/insect pests but are also great technical opportunities, for researchers, to silence gene expression (i.e. to study the effect of gene knock-out) in non genetically transformable obligate plant biotrophs, such as arbuscular mycorrhizal fungi (AMF).

In addition to HIGS and VIGS, a further RNAi-based approach to control plant pathogens has been proposed: the spray-induced gene silencing (SIGS). In this case, it has been observed that the simple application onto crops of sRNAs/dsRNAs targeting fungal genes (such as DCL, cytochrome P450, etc.) is sufficient to protect plants from a number of pathogens (Koch et al., 2016; Wang et al., 2016; McLoughlin et al., 2018). SIGS appears as a promising technique that could be likely used in agriculture in the next future (Huang et al., 2019).

All these findings, together, suggest that exchange of sRNAs between interacting organisms, and cross-kingdom RNAi-like processes, is likely an extremely widespread natural communication strategy.

1.7. The Arbuscular Mycorrhizal Symbiosis

The arbuscular mycorrhizal fungi (AMF) are key components of the plant soil microbiota which establish an intimate mutualistic association, known as arbuscular mycorrhizal symbiosis (AMS), with the majority of plant roots (Martin et al., 2017; Brundrett and Tedersoo, 2018). The key of the symbiosis is the reciprocal delivery of nutrients: AMF supplies plants with mineral nutrients, especially phosphate (Pi), in exchange of carbon compounds in the form of hexoses and lipids (Bravo et al., 2017; Jiang et al., 2017; Keymer et al., 2017; Luginbuehl et al., 2017; Ezawa and Saito, 2018).

Besides an improved mineral nutrition, mycorrhizal plants often show higher tolerance to abiotic and biotic stresses (Ruiz-Lozano, 2003; Göhre and Paszkowski, 2006; Liu et al., 2007; Jung et al., 2012; Chitarra et al., 2016; Miozzi et al., 2019). Furthermore, AMF colonization has an impact at systemic level and can influence several plant developmental processes such as flowering time, fruit and seed formation, and quality (Zouari et al., 2014; Fiorilli

et al., 2018). Due to these multiple benefits and the very wide host range, the relevance of AMF for applied purposes in the context of a more sustainable agriculture is therefore very high (Bender et al., 2016).

1.7.1. The fungal partner

AMF belong to the Glomeromycotina subphylum (Spatafora et al., 2016), a group of fungi characterized by peculiar features: - coenocytic hyphae and multinucleated asexual spores (Kamel et al., 2016; Lanfranco et al., 2016); - no sexual cycle, although mating-related genes are present in their genomes (Corradi and Brachmann, 2017); - obligate biotrophism, so uncultivable in the absence of plant hosts (Roth and Paszkowski, 2017); - large genomes extremely rich in transposable elements (Chen et al., 2018; Kobayashi et al., 2018; Morin et al., 2019; Sun et al., 2019; Venice et al., 2019). Furthermore, the lack of protocols for genome stable transformation makes the study of AMF, in a genetic perspective, extremely challenging.

Interestingly, genes encoding the cytosolic fatty acids (FA) synthase subunits, which are responsible for the bulk FA production in fungi, were not found in AMF genomes (Wewer et al., 2014; Tang et al., 2016). The characterization of plant mutant defective in genes involved in lipid metabolism clearly showed that AMF are entirely dependent on lipid supply by the plant for their growth, development and reproduction (Bravo et al., 2017; Jiang et al., 2017; Keymer et al., 2017; Luginbuehl et al., 2017). The dependence on host lipids may be the prime reason for their obligate biotrophy.

As many other fungi, AMF also possess their own microbiota (Bonfante et al., 2019), often hosting uncultivable endobacteria inside their cytoplasm (Bonfante and Desirò, 2017; Pawlowska et al., 2018). Molecular investigations, including the exploitation of transcriptomics data, have also allowed the description of viral sequences hosted within AMF (Turina et al., 2018). Mycoviruses can therefore be considered an additional component of the AMF microbiome with the potential to influence the biology of AMF and their host plant (Ikeda et al., 2012).

AMF, therefore, represent a fascinating case study for investigating the evolution of inter-kingdom interactions among plants, fungi, bacteria and viruses.

1.7.2. The colonization process

The establishment of a functional AMS requires a fine coordination at cellular and metabolic level of both partners; thus, a number of signaling molecules play a role in this inter-kingdom communication (MacLean et al., 2017; Lanfranco et al., 2018; Pimprikar and Gutjahr, 2018).

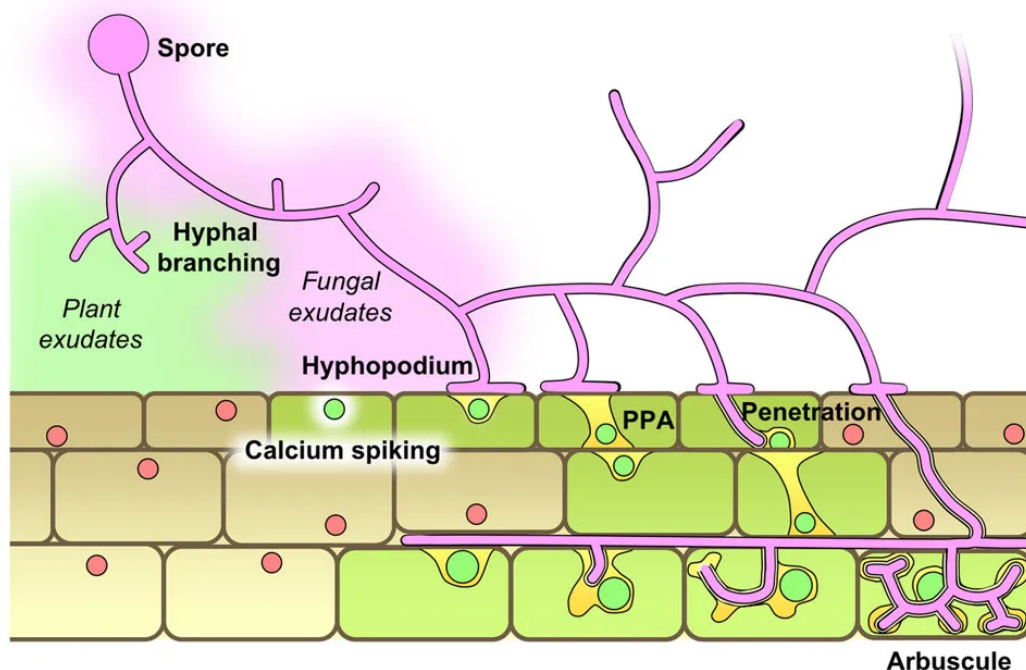


Figure 4. Schematic summary of the root colonization process by AM fungi (from Bonfante and Genre, 2010).

At the pre-symbiotic stage AMF recognize strigolactones (SL), plant hormones exuded by roots under low nutritional conditions. SL induce spore germination, hyphal branching, mitochondrial metabolism and the release by AMF of chitin oligosaccharides (Myc-COs) and lipochitooligosaccharides (Myc-LCOs), known as “Myc factors” (Fig. 4). The Myc Factors are perceived by the root epidermis where they trigger a calcium spiking through the activation of the so-called common symbiosis signaling pathway (the same involved in nodulation; Oldroyd, 2013; MacLean et al., 2017) that prepare the plant for fungal accommodation and induce the expression of a range of

symbiosis-supporting genes (Schmitz and Harrison, 2014; Hohnjec et al., 2015). These include transcription factors which coordinate the expression of a set of downstream genes involved in specific symbiotic functions, such as phosphate transport and lipid synthesis and transport (Pimprikar and Gutjahr, 2018). Root colonization starts with the differentiation of fungal hyphopodia which allow the penetration of the plant epidermis (Bonfante and Genre, 2010; Fig. 4). The fungal penetration is favored by a specific plant sub-cellular rearrangement, known as pre-penetration apparatus (PPA), that consists of a cytoplasmic bridge across the vacuole of epidermal and underlying outer cortical cells (Genre et al., 2005). The fungal hyphae develop into the PPA until they reach the inner cortex where, at this point, they grow in the intercellular space along the root axis. Finally, PPA is also constituted during the fungal entry of cortical cells, where after an extensive dichotomous branching fungal hyphae form arbuscules (Genre et al., 2008; Fig. 4, 5).

Arbuscules are always surrounded by the plant plasma membrane, known as periarbuscular membrane, that constitutes, together with the fungal plasma membrane, the final interface for nutrient exchange between AMF and plants (Luginbuehl and Oldroyd, 2017; Fig. 5). The space between plant and AMF plasma membranes inside arbusculated cells is known as periarbuscular space. Recently, extensive membranes rearrangements (such as membrane tubules) and extracellular vesicles formation have been observed in the periarbuscular space (Ivanov et al., 2019; Roth et al. 2019). It has been hypothesized that they may be involved in the exchange of signal molecules and/or nutrients between the partners.

The functional life time of arbuscules has been estimated a few days (Kobae and Hata, 2010) after which senescence processes are initiated: progressively larger branches and the arbuscules trunk septate, collapse and eventually disappear while the host cell remains active and maintains the ability to be colonized again by a new arbuscule (Luginbuehl and Oldroyd, 2017). One candidate marker for arbuscule senescence is the plant MYB1 transcription factor which is required for the expression of several *Medicago truncatula* genes, including hydrolases, associated with arbuscule collapse (Floss et al., 2017).

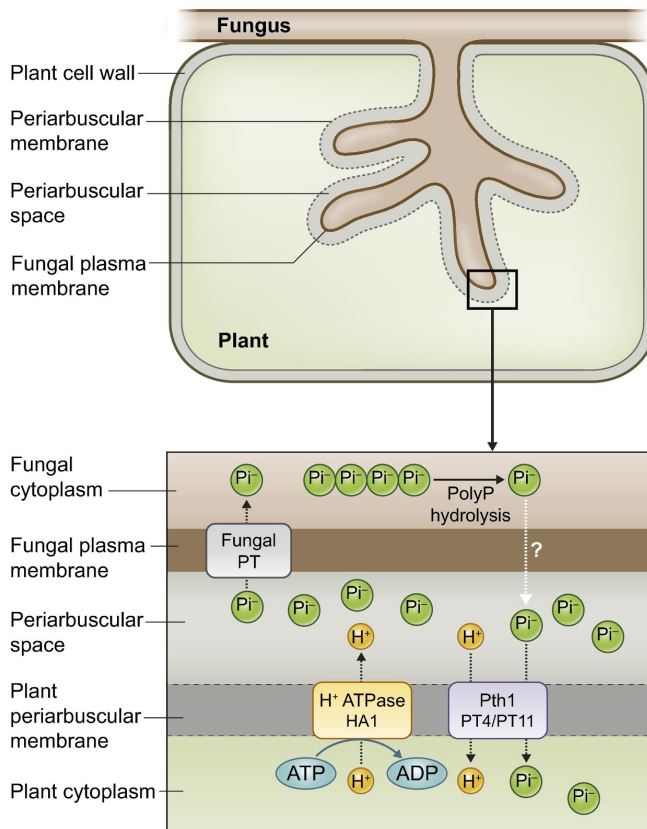


Figure 5. Schematic representation of arbuscule structure (top) with details about phosphate transfer mechanism from AMF to host cell (bottom; from Lanfranco et al., 2018).

1.7.3. Flux and role of nutrients (phosphorus and carbon) in AMS

The Pi transport in AMS, a key feature of the symbiosis, has been extensively studied (Ezawa and Saito, 2018). In extraradical mycelium, Pi collected from soil thanks to ATPase-dependent H⁺/Pi symporters (PT), is quickly converted inside vacuoles into polyphosphate (polyP) chains. Kikuchi et al. (2016) proposed a model in which transpiration provides a primary driving force for polyP

translocation by creating water flow through the fungal RcAQP3 and the arbuscular mycorrhizal (AM)-inducible plant aquaporins. Once in the arbuscules, polyP are hydrolyzed, possibly thanks to acid and alkaline phosphatases, and then Pi is exported in the apoplast through a still unknown

process, which may involve the fungal plasma-membrane transporters or Golgi/trans-Golgi network-mediated secretion pathway (Ezawa and Saito, 2018). From the periarbuscular space Pi is taken up and transferred into plant cell cytoplasm by PT, belonging to the Phosphate transporter 1 (Pht1) class, localized at the plant periarbuscular membrane (Harrison et al., 2002; Javot et al., 2007; Luginbuehl and Oldroyd, 2017; Fig. 5); this transport is suggested to be driven by an H⁺ energy gradient produced by an H⁺-ATPase (Krajinski et al., 2014; Wang et al., 2014; Fig. 5). The expression of genes encoding for these PTs - which have been characterized in several AM-host plants - is specifically induced by AMF root colonization in arbusculated cells. For this reason, these PTs, including the *Medicago truncatula* MtPT4, are exploited as molecular markers for arbuscule-containing cells and for a functional AMS.

It is also worth noting that the fungal PT genes responsible of the uptake of Pi from the soil solutions are expressed in the extraradical mycelium but also in arbusculated cells (Balestrini et al., 2007; Fig. 5) suggesting a second role in Pi reabsorption from the periarbuscular space.

Interestingly, PT4 promoters of *M. truncatula* and *Lotus japonicus* are also active in root tips when grown in Pi starvation conditions; the corresponding mutant lines do not fully respond to low Pi with changes in lateral root formation (Volpe et al., 2016), suggesting that PT4 are involved not only in symbiotic Pi uptake but also in Pi sensing controlling root architecture responses to low Pi. A role in Pi sensing has also been proposed for the PT expressed in arbusculated cells (Lanfranco et al., 2018 and references therein). Pi does not have only a nutritional role as the Pi status influences AMS development: when a fungal PT or plant PT genes essential for AMS are silenced, arbuscule development is affected (Javot et al., 2007; Volpe et al., 2016; Xie et al., 2016). In addition to its cell-autonomous influence on arbuscule maintenance, Pi also regulates AMS in a systemic manner as AM colonization is repressed when plants are grown under high Pi supply; the mechanisms underlying this effect are not known (Lanfranco et al., 2018).

AMF receive carbohydrates as well as lipids from the host. Stable isotope labelling experiments showed that AMF receive glucose from the plant (Pfeffer et al., 1999). While a number of genes encoding sugar transporters with activities towards monosaccharides (MSTs) and sucrose (SUTs), and members of the SWEET family, were shown to be up-regulated in mycorrhizal roots

(Harrison, 1996; Doidy et al., 2012; Manck-Götzenberger and Requena, 2016), so far their specific involvement in C transfer was not demonstrated. On the fungal side, a high affinity monosaccharide transporter MST2 from the AMF *R. irregularis*, expressed in arbuscules and intercellular hyphae, was hypothesized to be responsible for fungal sugar uptake (Helber et al., 2011).

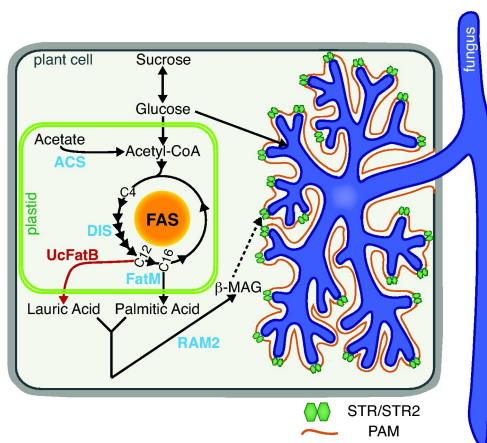


Figure 6. Model of the AM-specific lipid biosynthesis pathway in a plant cell hosting an arbuscule (from Keymer and Gutjahr, 2018). The red arrow illustrates the synthetic deviation of the pathway by transgenic expression of *UcFatB* in *Medicago* hairy roots. ACS, acetyl-coenzym A synthase; DIS, DISORGANIZED ARBUSCULES; *FatM*, Mycorrhizal acyl-ACP thioesterase; RAM2, REDUCED ARBUSCULAR MYCORRHIZA 2; STR, STUNTED ARBUSCULE; PAM, peri-arbuscular membrane.

An important discovery of the last years, that may also explain the AMF obligate biotrophy, is that lipids are transferred from the host plant to AMF which are indeed auxotrophs for fatty acids (FA) as they lack genes encoding fatty acid synthase I subunits (Bravo et al., 2017; Jiang et al., 2017; Keymer et al., 2017; Luginbuehl et al., 2017). Plant mutant defective in three AM-induced lipid biosynthetic genes: *DISORGANIZED ARBUSCULES* (*DIS*, encoding a keto-acyl ACP synthase I), *FatM* (encoding an acyl-ACP-acyl carrier protein-thioesterase) and *REDUCED ARBUSCULAR MYCORRHIZA 2* (*RAM2*, encoding a glycerol-3-phosphate acyltransferase) showed severe defects in AMF colonization and arbuscule morphology (Wang et al., 2012; Bravo et al., 2017; Jiang et al., 2017; Keymer et al., 2017; Luginbuehl et al., 2017). The flux of lipids from the host plant to AMF was elegantly demonstrated by different experimental approaches (for a review see Keymer and Gutjahr, 2018). Some

aspects of the flux of lipids remain to be defined but from these studies it has been suggested that a candidate-lipid to be exported from the plant cell may be 16:0-MAG (monoacylglycerol) and a candidate to pump lipids from the plant cell into the periarbuscular space, could be the ABCG transporter complex, consisting of two half transporters STR and STR2, which is also specifically expressed in arbusculated cells (Zhang et al., 2010; Gutjahr et al., 2012; Fig. 6).

1.7.4. An emerging role for fungal effectors, peptides and small RNA in AMS

The establishment of AMS is likely to be mediated by AMF effectors, in analogy to other plant-microbe interactions. A large number of effectors have been predicted from AMF genomes and transcriptomes (Kamel et al., 2017) by looking at the presence of a signal peptide that guides proteins towards secretion. So far, the only functional characterized AMF effector is the protein named SP7 from *Rhizophagus irregularis*; SP7 was suggested to counteract the plant immune response by interacting in plant cell nuclei with the pathogenesis-related-transcription factor ethylene response factor ERF19 (Kloppholz et al., 2011). Another AMF effector (RiCRN1) was also recently shown to localize in plant nuclei and to be critical for symbiosis progression and arbuscule development, although the mechanism of action is not known (Voß et al., 2018). Zeng et al. (2019) added an important element to this scenario, revealing that a fungal secreted proteins (RiSLM), containing a LysM motif and showing high affinity for long chain chitin molecules, can play a role in sequestering these molecules that are known to act as plant defence elicitors; thus, RiSLM could promote the recognition of symbiotic signals in AMF exudates.

Recently, a new class of fungal-secreted regulators of plant AMS-related genes has been discovered: the CLAVATA/EMBRYO SURROUNDING REGION (ESR)-RELATED PROTEIN-like (CLE-like; Le Marquer et al., 2019). AMF can indeed produce and secrete CLE-like molecules, that may possibly mimic the plant CLE peptides, a group of endogenous regulators of different proliferation/differentiation processes (Czyzewicz et al., 2013). Plant treatment with the *R. irregularis* CLE-like peptide RiCLE1 induces the formation of lateral roots and inhibits primary root growth (Le Marquer et al., 2019), compatible

with a role of this peptide in the previous observed change of root architecture in mycorrhizal plants (Fusconi, 2014); in addition, RiCLE1 also stimulates mycorrhization (Le Marquer et al., 2019). The same authors also showed that the host plant *M. truncatula* produces CLE peptides (MtCLE16 and MtCLE52) nearly identical to RiCLE1, whose genes are up-regulated during mycorrhization (Le Marquer et al., 2019). Interestingly, another *M. truncatula* CLE peptide (MtCLE53) was recently characterized and found to negatively regulate AM colonization levels by modulating root strigolactones content. These findings on plant CLE and fungal CLE-like peptides add a new layer of complexity to the full understanding of partner communication in AMS mediated by peptides/proteins.

MicroRNAs could also participate in the AMS development and functioning (reviewed in Lelandais-Brière et al., 2016; Müller and Harrison, 2019). Members of the miR399 family, which are systemic Pi-starvation signals, have been proposed as signalling molecules in the regulation of AMS by Pi, as they are induced in mycorrhizal roots (Branscheid et al., 2010). However, miR399 overexpression did not restore AM fungal colonization at high Pi concentration (Branscheid et al., 2010), suggesting that other mechanisms are involved.

The expression of some plant AMS-related genes is under control of miRNAs (Muller and Harrison 2019). The best characterized example concerns miRNAs of the conserved family miR171, known for regulating the expression of *GRAS transcription factors (TFs)*; Cenci and Rouard, 2017). Notably, in *M. truncatula* it was reported that the isoform *mtr-miR171h*, targeting the *NSP2 (Nodulation Signaling Pathway 2; a GRAS TF)*, is specifically induced by AMF or Myc-LCO treatments (Branscheid et al., 2011; Devers et al., 2011; Lauressergues et al., 2012; Hofferek et al., 2014). In plant, *NSP2* cooperates with other *GRAS TFs*, such as *NSP1* and *RAM1 (Required For Arbuscular Mycorrhization-1)*, to promote the expression of AMS-specific genes, that are required for correct AMF root colonization (Gobbato et al., 2012; Rich et al., 2015). In particular, *mtr-miR171h* is specifically expressed in those plant peripheral tissues (e.g. epidermis and first cellular layer of the cortex of the elongation zone) in which the colonization or over-colonization by AMF should be prevented. Furthermore, the miR171 family is also known for controlling the expression of *LOM (Lost Meristems)* genes, another group of *GRAS TFs* (Xue et al., 2014); in *M. truncatula* *LOM1* is a positive regulator of AMS (Couzigou et al., 2017).

Interestingly, the isoform *mtr*-miR171b, that is encoded only by AMS-competent plants, is specifically induced in cells containing arbuscules and protects *LOM1* from being targeted by other members of the miR171 family, thanks to a mismatch in the cleavage site (Couzigou et al., 2017). Indeed, the mismatch makes AGO-miR171b complex cleavage-incompetent towards *LOM1* transcripts and, competing for the same substrate with other cleavage-competent miR171s, miR171b prevents the silencing of its target.

AMF colonization is also positively regulated in plant by auxin level (Liao et al., 2018). From a molecular point of view, auxin triggers the ubiquitination mediated by SCF^{TIR1/AFB} (and thus degradation by 26S proteasome) of the AUX/IAA proteins which are constitutive inhibitors of the *ARF TFs*, the final inducers of the transcription of the auxin-responsive genes (Salehin et al., 2015). In this negative regulatory circuit, miR393 participates modulating the expression, at PTGS level, of the *TIR1/AFB* components of the SCF^{TIR1/AFB} complex; thus, miR393 is a negative regulator of auxin response (Vidal et al., 2010). It has been shown that miR393 is down-regulated in mycorrhizal roots, suggesting that TIR1/AFB-dependent auxin signaling is required for arbuscule formation (Etemadi et al., 2014).

Furthermore, it has been shown that plant miR396 controls mycorrhizal response by targeting *GRF (Growth Regulation Factor)* genes that are important regulators of cell proliferation (Bazin et al., 2013).

Besides the previous examples, in which the molecular details of plant miRNA-mediated gene regulation has been characterized, large scale omics screenings revealed that AMS impacts the expression of several other plant miRNAs, dozens of which are conserved between different plant species (Gu et al., 2010; Devers et al., 2011; Formey et al., 2014; Wu et al., 2016; Pandey et al., 2018). In *M. truncatula*, most of the predicted miRNA targets upon AMF colonization are represented by transcription factors and disease resistant genes (Devers et al., 2011). Partially similar results were reported in *Nicotiana attenuata*: in this case, most of AMS-related miRNA targets seem to be involved in defense, Pi starvation and phytohormone signaling (Pandey et al., 2011). It is worth noting that, so far, all the knowledge about “sRNA world” in AMS was available only for plant species.

1.8. Aims of the PhD thesis

Small RNAs are nowadays recognized as key players in several inter-kingdom interactions, through cross-kingdom RNAi-like processes. Considering the intimate nature of the interaction between plant and AMF, the main long-term aim of this research line is to understand whether silencing-competent sRNAs are exchanged between the symbiotic partners of AMS, in particular moving from the fungus to the host plant. However, at the beginning of my PhD, nothing was known about RNAi and sRNAs in AMF. For this reason during my PhD, I have first characterized the RNAi machinery and the small RNA population in AMF, with a specific focus on their potential role in host gene regulation. I have also contributed to the first characterization of the virome of an AMF (*Gigaspora margarita*) and provided evidence that its RNAi may function as an antiviral defense.

In more detail, during my PhD I developed these specific research activities:

1. Characterization of the RNAi machinery and the smallRNAome in the model AMF *Rhizophagus irregularis* with a focus on their modulation in two fungal life stages, the extra-radical and intra-radical mycelium. I also investigated *in silico* evidence of cross-kingdom RNAi in the association between this AMF with its host plant *M. truncatula*. This work, which is presented in **Chapter 2**, was published by *BMC Genomics* in 2019
2. Experimental validation of some of the *in silico* prediction presented in Chapter 2. In particular, I evaluated the effectiveness of *in silico* target predictions by means of co-expression assays in tobacco leaves; this work was carried out in collaboration with Prof. Hailing Jin's Lab at University of California, Riverside (UCR) where I spent 6 months. I also investigated the expression profiles of the identified plant targets of AMF sRNAs in specific root cell populations by means of qRT-PCR coupled to Laser Microdissection (LMD). The results are shown in **Chapter 3**.
3. Characterization of the viral population infecting the AMF *Gigaspora margarita*. In particular, my contribution was on the identification and

characterization of DNA fragments, not integrated into mitochondrial nor nuclear genome, that correspond to the RNA genomes of 4 mitoviruses. This work was published in *Environmental Microbiology* in 2018 and it is here presented in **Chapter 4**.

4. Description of the RNAi machinery and sRNA populations in *G. margarita*. I performed a comparative analysis with the findings obtained on *R. irregularis* (Chapter 2) and I determined all the potential genetic sources that can originate sRNAs (nuclear, mitochondrial, viral and endobacterial genomes). I also reported *in silico* evidence of a possible role of viral-derived sRNAs in cross-kingdom RNAi. This work, presented in **Chapter 5**, is under revision on *Frontiers in Microbiology*.

1.9. References

- Axtell, M. J. (2013). Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol.* 64, 137–159. doi:10.1146/annurev-arplant-050312-120043.
- Axtell, M. J., and Meyers, B. C. (2018). Revisiting criteria for plant microRNA annotation in the Era of big data. *Plant Cell* 30, 272–284. doi:10.1105/tpc.17.00851.
- Balestrini, R., Gómez-Ariza, J., Lanfranco, L., and Bonfante, P. (2007). Laser microdissection reveals that transcripts for five plant and one fungal phosphate transporter genes are contemporaneously present in arbusculated cells. *Mol. Plant-Microbe Interact.* 20, 1055–1062. doi:10.1094/MPMI-20-9-1055.
- Bao, X., and Roossinck, M. J. (2013). Multiplexed Interactions: viruses of endophytic Fungi. *Adv. Virus Res.* 86, 37–58. doi:10.1016/B978-0-12-394315-6.00002-7.
- Baulcombe, D. C. (2015). VIGS, HIGS and FIGS: small RNA silencing in the interactions of viruses or filamentous organisms with their plant hosts. *Curr. Opin. Plant Biol.* 26, 141–146. doi:10.1016/j.pbi.2015.06.007.
- Bazin, J., Khan, G. A., Combier, J. P., Bustos-Sanmamed, P., Debernardi, J. M., Rodriguez, R., et al. (2013). MiR396 affects mycorrhization and root meristem activity in the legume *Medicago truncatula*. *Plant J.* 74, 920–934. doi:10.1111/tpj.12178.
- Bender, S. F., Wagg, C., and van der Heijden, M. G. A. (2016). An underground revolution: biodiversity and soil ecological engineering for agricultural sustainability. *Trends Ecol. Evol.* 31, 440–452. doi:10.1016/j.tree.2016.02.016.
- Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363–366. doi:10.1038/35053110.
- Bonfante, P., and Desirò, A. (2017). Who lives in a fungus? The diversity, origins and functions of fungal endobacteria living in Mucoromycota. *ISME J.* 11, 1727–1735. doi:10.1038/ismej.2017.21.

- Bonfante, P., and Genre, A. (2010). Mechanisms underlying beneficial plant–fungus interactions in mycorrhizal symbiosis. *Nat. Commun.* 1, 48. doi:10.1038/ncomms1046.
- Bonfante, P., Venice, F., and Lanfranco, L. (2019). The mycobiota: fungi take their place between plants and bacteria. *Curr. Opin. Microbiol.* 49, 18–25. doi:10.1016/j.mib.2019.08.004.
- Borges, F., and Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nat. Publ. Gr.* 16, 1–15. doi:10.1038/nrm4085.
- Branscheid, A., Devers, E. A., May, P., and Krajinski, F. (2011). Distribution pattern of small RNA and degradome reads provides information on miRNA gene structure and regulation. *Plant Signal. Behav.* 6, 1609–1611. doi:10.4161/psb.6.10.17305.
- Branscheid, A., Sieh, D., Datt Pant, B., May, P., Devers, E. A., Elkrog, A., et al. (2010). Expression pattern suggests a role of MiR399 in the regulation of the cellular response to local Pi increase during arbuscular mycorrhizal symbiosis. *Mol. Plant-Microbe Interact.* 23, 915–926. doi:10.1094/MPMI-23-7-0915.
- Bravo, A., Brands, M., Wewer, V., Dörmann, P., and Harrison, M. J. (2017). Arbuscular mycorrhiza-specific enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular mycorrhiza. *New Phytol.* 214, 1631–1645. doi:10.1111/nph.14533.
- Brundrett, M. C., and Tedersoo, L. (2018). Evolutionary history of mycorrhizal symbioses and global host plant diversity. *New Phytol.* 220, 1108–1115. doi:10.1111/nph.14976.
- Buck, A. H., Coakley, G., Simbari, F., McSorley, H. J., Quintana, J. F., Le Bihan, T., et al. (2014). Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity. *Nat. Commun.* 5, 5488. doi:10.1038/ncomms6488.
- Cai, Q., Qiao, L., Wang, M., He, B., Lin, F.-M., and Palmquist, J. (2018). Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. *Science.* doi:10.1126/science.aar4142.
- Calo, S., Nicolás, F. E., Vila, A., Torres-Martínez, S., and Ruiz-Vázquez, R. M. (2012). Two distinct RNA-dependent RNA polymerases are required for initiation and amplification of RNA silencing in the basal fungus *Mucor circinelloides*. *Mol. Microbiol.* 83, 379–394. doi:10.1111/j.1365-2958.2011.07939.x.
- Campo, S., Gilbert, K. B., and Carrington, J. C. (2016). Small RNA-based antiviral defense in the phytopathogenic fungus *Colletotrichum higginsianum*. *PLOS Pathog.* 12, e1005640. doi:10.1371/journal.ppat.1005640.
- Cenci, A., and Rouard, M. (2017). Evolutionary analyses of GRAS transcription factors in angiosperms. *Front. Plant Sci.* 8, 273. doi:10.3389/fpls.2017.00273.
- Cenik, E. S., and Zamore, P. D. (2011). Argonaute proteins. *Curr. Biol.* 21, R446–R449. doi:10.1016/j.cub.2011.05.020.
- Chaloner, T., van Kan, J. A. L., and Grant-Downton, R. T. (2016). RNA ‘information warfare’ in pathogenic and mutualistic interactions. *Trends Plant Sci.* 21, 738–748. doi:10.1016/j.tplants.2016.05.008.
- Chen, E. C. H., Morin, E., Beaudet, D., Noel, J., Yildirim, G., Ndikumana, S., et al. (2018). High intraspecific genome diversity in the model arbuscular mycorrhizal symbiont *Rhizophagus irregularis*. *New Phytol.* doi:10.1111/nph.14989.

- Chitarra, W., Maserti, B., Gambino, G., Guerrieri, E., and Balestrini, R. (2016). Arbuscular mycorrhizal symbiosis-mediated tomato tolerance to drought. *Plant Signal. Behav.* 11, e1197468. doi:10.1080/15592324.2016.1197468.
- Chow, F. W. N., Koutsovoulos, G., Ovando-Vázquez, C., Neophytou, K., Bermúdez-Barrientos, J. R., Laetsch, D. R., et al. (2019). Secretion of an Argonaute protein by a parasitic nematode and the evolution of its siRNA guides. *Nucleic Acids Res.* 47, 3594–3606. doi:10.1093/nar/gkz142.
- Christie, M., Croft, L. J., and Carroll, B. J. (2011). Intron splicing suppresses RNA silencing in *Arabidopsis*. *Plant J.* 68, 159–167. doi:10.1111/j.1365-313X.2011.04676.x.
- Corradi, N., and Brachmann, A. (2017). Fungal mating in the most widespread plant symbionts? *Trends Plant Sci.* 22, 175–183. doi:10.1016/j.tplants.2016.10.010.
- Couzigou, J. M., Laressergues, D., André, O., Gutjahr, C., Guillotin, B., Bécard, G., et al. (2017). Positive gene regulation by a natural protective miRNA enables arbuscular mycorrhizal symbiosis. *Cell Host Microbe* 21, 106–112. doi:10.1016/j.chom.2016.12.001.
- Cui, C., Wang, Y., Liu, J., Zhao, J., Sun, P., and Wang, S. (2019). A fungal pathogen deploys a small silencing RNA that attenuates mosquito immunity and facilitates infection. *Nat. Commun.* 10, 4298. doi:10.1038/s41467-019-12323-1.
- Cuperus, J. T., Fahlgren, N., and Carrington, J. C. (2011). Evolution and functional diversification of MIRNA genes. *Plant Cell* 23, 431–442. doi:10.1105/tpc.110.082784.
- Czyzewicz, N., Yue, K., Beeckman, T., and De Smet, I. (2013). Message in a bottle: small signalling peptide outputs during growth and development. *J. Exp. Bot.* 64, 5281–5296. doi:10.1093/jxb/ert283.
- Dang, Y., Li, L., Guo, W., Xue, Z., and Liu, Y. (2013). Convergent transcription induces dynamic DNA methylation at disiRNA loci. *PLoS Genet.* 9, e1003761. doi:10.1371/journal.pgen.1003761.
- Dang, Y., Zhang, Z., and Liu, Y. (2014). Small RNA-mediated gene silencing in *Neurospora*. *Fungal RNA Biology*, 1–395. doi:10.1007/978-3-319-05687-6.
- Devers, E. A., Branscheid, A., May, P., and Krajinski, F. (2011). Stars and symbiosis: microRNA- and microRNA*-mediated transcript cleavage involved in arbuscular mycorrhizal symbiosis. *Plant Physiol.* 156, 1990–2010. doi:10.1104/pp.111.172627.
- Doidy, J., Van Tuinen, D., Lamotte, O., Corneillat, M., Alcaraz, G., and Wipf, D. (2012). The *Medicago truncatula* sucrose transporter family: characterization and implication of key members in carbon partitioning towards arbuscular mycorrhizal fungi. *Mol. Plant* 5, 1346–1358. doi:10.1093/mp/sss079.
- Donaldson, M. E., and Saville, B. J. (2012). Natural antisense transcripts in fungi. *Mol. Microbiol.* 85, 405–417. doi:10.1111/j.1365-2958.2012.08125.x.
- Drinnenberg, I. A., Fink, G. R., and Bartel, D. P. (2011). Compatibility with killer explains the rise of RNAi-deficient fungi. *Science* 333, 1592. doi:10.1126/science.1209575.
- Etemadi, M., Gutjahr, C., Couzigou, J.-M., Zouine, M., Laressergues, D., Timmers, A., et al. (2014). Auxin perception is required for arbuscule development in arbuscular mycorrhizal symbiosis. *Plant Physiol.* 166, 281–92. doi:10.1104/pp.114.246595.

- Ezawa, T., and Saito, K. (2018). How do arbuscular mycorrhizal fungi handle phosphate? New insight into fine-tuning of phosphate metabolism. *New Phytol.* 220, 1116–1121. doi:10.1111/nph.15187.
- Fei, Q., Xia, R., and Meyers, B. C. (2013). Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell* 25, 2400–15. doi:10.1105/tpc.113.114652.
- Feretzi, M., Billmyre, R. B., Clancey, S. A., Wang, X., and Heitman, J. (2016). Gene network polymorphism illuminates loss and retention of novel RNAi silencing components in the *Cryptococcus* pathogenic species complex. *PLOS Genet.* 12, e1005868. doi:10.1371/journal.pgen.1005868.
- Fiorilli, V., Vannini, C., Ortolani, F., Garcia-Secco, D., Chiapello, M., Novero, M., et al. (2018). Omics approaches revealed how arbuscular mycorrhizal symbiosis enhances yield and resistance to leaf pathogen in wheat. *Sci. Rep.* 8, 9625. doi:10.1038/s41598-018-27622-8.
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811.
- Floss, D. S., Gomez, S. K., Park, H. J., MacLean, A. M., Müller, L. M., Bhattarai, K. K., et al. (2017). A Transcriptional program for arbuscule degeneration during AM symbiosis is regulated by MYB1. *Curr. Biol.* 27, 1206–1212. doi:10.1016/j.cub.2017.03.003.
- Formey, D., Sallet, E., Lelandais-Brière, C., Ben, C., Bustos-Sanmamed, P., Niebel, A., et al. (2014). The small RNA diversity from *Medicago truncatula* roots under biotic interactions evidences the environmental plasticity of the miRNAome. *Genome Biol.* 15, 457. doi:10.1186/s13059-014-0457-4.
- Fusconi, A. (2014). Regulation of root morphogenesis in arbuscular mycorrhizae: What role do fungal exudates, phosphate, sugars and hormones play in lateral root formation? *Ann. Bot.* 113, 19–33. doi:10.1093/aob/mct258.
- Genre, A., Chabaud, M., Faccio, A., Barker, D. G., and Bonfante, P. (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, 1407–1420. doi:10.1105/tpc.108.059014.
- Genre, A., Chabaud, M., Timmers, T., Bonfante, P., and Barker, D. G. (2005). Arbuscular mycorrhizal fungi elicit a novel intracellular apparatus in *Medicago truncatula* root epidermal cells before infection. *Plant Cell* 17, 3489–3499. doi:10.1105/tpc.105.035410.
- Ghabrial, S. A., Castón, J. R., Jiang, D., Nibert, M. L., and Suzuki, N. (2015). 50-plus years of fungal viruses. *Virology* 479–480, 356–368. doi:10.1016/J.VIROL.2015.02.034.
- Ghildiyal, M., and Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* 10, 94–108. doi:10.1038/nrg2504.
- Gobbato, E., Marsh, J. F., Vernié, T., Wang, E., Maillet, F., Kim, J., et al. (2012). A GRAS-type transcription factor with a specific function in mycorrhizal signaling. *Curr. Biol.* 22, 2236–2241. doi:10.1016/j.cub.2012.09.044.
- Göhre, V., and Paszkowski, U. (2006). Contribution of the arbuscular mycorrhizal symbiosis to heavy metal phytoremediation. *Planta* 223, 1115–1122. doi:10.1007/s00425-006-0225-0.

- Gu, M., Xu, K., Chen, A., Zhu, Y., Tang, G., and Xu, G. (2010). Expression analysis suggests potential roles of microRNAs for phosphate and arbuscular mycorrhizal signaling in *Solanum lycopersicum*. *Physiol. Plant.* 138, 226–237. doi:10.1111/j.1399-3054.2009.01320.x.
- Gutjahr, C., Radovanovic, D., Geoffroy, J., Zhang, Q., Siegler, H., Chiapello, M., et al. (2012). The half-size ABC transporters STR1 and STR2 are indispensable for mycorrhizal arbuscule formation in rice. *Plant J.* 69, 906–920. doi:10.1111/j.1365-313X.2011.04842.x.
- Hammond, T. M., Andrewski, M. D., Roossinck, M. J., and Keller, N. P. (2008). *Aspergillus* mycoviruses are targets and suppressors of RNA silencing. *Eukaryot. Cell* 7, 350–7. doi:10.1128/EC.00356-07.
- Harrison, M. J. (1996). A sugar transporter from *Medicago truncatula*: altered expression pattern in roots during vesicular-arbuscular (VA) mycorrhizal associations. *Plant J.* 9, 491–503. doi:10.1046/j.1365-313x.1996.09040491.x.
- Harrison, M. J., Dewbre, G. R., and Liu, J. (2002). A phosphate transporter from *Medicago truncatula* involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. *Plant Cell* 14, 2413–29. doi:10.1105/TPC.004861.
- Helber, N., Wippel, K., Sauer, N., Schaarschmidt, S., Hause, B., and Requena, N. (2011). A Versatile monosaccharide transporter that operates in the arbuscular mycorrhizal fungus *Glomus* sp is crucial for the symbiotic relationship with plants. *Plant Cell* 23, 3812–3823. doi:10.1105/tpc.111.089813.
- Henz, S. R., Cumbie, J. S., Kasschau, K. D., Lohmann, J. U., Carrington, J. C., Weigel, D., et al. (2007). Distinct expression patterns of natural antisense transcripts in *Arabidopsis*. *Plant Physiol.* 144, 1247–1255. doi:10.1104/PP.107.100396.
- Herrero, N., Sánchez Márquez, S., and Zabalgoceazcoa, I. (2009). Mycoviruses are common among different species of endophytic fungi of grasses. *Arch. Virol.* 154, 327–330. doi:10.1007/s00705-008-0293-5.
- Hofferek, V., Mendrinna, A., Gaude, N., Krajinski, F., and Devers, E. A. (2014). MiR171h restricts root symbioses and shows like its target NSP2 a complex transcriptional regulation in *Medicago truncatula*. *BMC Plant Biol.* 14, 199. doi:10.1186/s12870-014-0199-1.
- Hohnjec, N., Czaja-Hasse, L. F., Hogeckamp, C., and Küster, H. (2015). Pre-announcement of symbiotic guests: transcriptional reprogramming by mycorrhizal lipochitooligosaccharides shows a strict co-dependency on the GRAS transcription factors NSP1 and RAM1. *BMC Genomics* 16, 994. doi:10.1186/s12864-015-2224-7.
- Hou, Y., Zhai, Y., Feng, L., Karimi, H. Z., Rutter, B. D., Zeng, L., et al. (2019). A *Phytophthora* effector suppresses trans-kingdom RNAi to promote disease susceptibility. *Cell Host Microbe* 25, 153–165.e5. doi:10.1016/j.chom.2018.11.007.
- Huang, C.-Y., Wang, H., Hu, P., Hamby, R., and Jin, H. (2019). Small RNAs – big players in plant-microbe interactions. *Cell Host Microbe* 26, 173–182. doi:10.1016/j.chom.2019.07.021.
- Hutvagner, G., and Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* 9, 22–32. doi:10.1038/nrm2321.
- Ikeda, Y., Shimura, H., Kitahara, R., Masuta, C., and Ezawa, T. (2012). A novel virus-like double-stranded RNA in an obligate biotroph arbuscular mycorrhizal fungus: a hidden

- player in mycorrhizal symbiosis. *Mol. Plant-Microbe Interact.* 25, 1005–1012. doi:10.1094/MPMI-11-11-0288.
- Ipsaro, J. J., and Joshua-Tor, L. (2015). From guide to target: molecular insights into eukaryotic RNA-interference machinery. *Nat. Struct. Mol. Biol.* 22, 20–28. doi:10.1038/nsmb.2931.
- Ivanov, S., Austin, J., Berg, R. H., and Harrison, M. J. (2019). Extensive membrane systems at the host–arbuscular mycorrhizal fungus interface. *Nat. Plants* 5, 194–203. doi:10.1038/s41477-019-0364-5.
- Javot, H., Penmetsa, R. V., Terzaghi, N., Cook, D. R., and Harrison, M. J. (2007). A *Medicago truncatula* phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1720–1725. doi:10.1073/pnas.0608136104.
- Jiang, Y., Wang, W., Xie, Q., Liu, N., Wang, D., Zhang, X., et al. (2017). Plants transfer lipids to sustain colonization by mutualistic mycorrhizal and parasitic fungi. *Science*. 356, 1172–1175. doi:10.1126/science.aam9970.
- Jiao, J., and Peng, D. (2018). Wheat microRNA1023 suppresses invasion of *Fusarium graminearum* via targeting and silencing *FGSG_03101*. *J. Plant Interact.* 13, 514–521. doi:10.1080/17429145.2018.1528512.
- Jones-Rhoades, M. W. (2012). Conservation and divergence in plant microRNAs. *Plant Mol. Biol.* 80, 3–16. doi:10.1007/s11103-011-9829-2.
- Juliano, C., Wang, J., and Lin, H. (2011). Uniting germline and stem cells: the function of Piwi proteins and the piRNA pathway in diverse organisms. *Annu. Rev. Genet.* 45, 447–469. doi:10.1146/annurev-genet-110410-132541.
- Jung, S. C., Martinez-Medina, A., Lopez-Raez, J. A., and Pozo, M. J. (2012). Mycorrhiza-induced resistance and priming of plant defenses. *J. Chem. Ecol.* 38, 651–664. doi:10.1007/s10886-012-0134-6.
- Kamel, L., Keller-Pearson, M., Roux, C., and Ané, J.-M. (2016). Biology and evolution of arbuscular mycorrhizal symbiosis in the light of genomics. *New Phytol.*, 57–64.
- Kamel, L., Tang, N., Malbreil, M., San Clemente, H., Le Marquer, M., Roux, C., et al. (2017). The comparison of expressed candidate secreted proteins from two arbuscular mycorrhizal fungi unravels common and specific molecular tools to invade different host plants. *Front. Plant Sci.* 8, 124. doi:10.3389/fpls.2017.00124.
- Keymer, A., and Gutjahr, C. (2018). Cross-kingdom lipid transfer in arbuscular mycorrhiza symbiosis and beyond. *Curr. Opin. Plant Biol.* 44, 137–144. doi:10.1016/j.pbi.2018.04.005.
- Keymer, A., Pimprikar, P., Wewer, V., Huber, C., Brands, M., Bucerius, S. L., et al. (2017). Lipid transfer from plants to arbuscular mycorrhiza fungi. *Elife* 6, e29107. doi:10.7554/eLife.29107.
- Kikuchi, Y., Hijikata, N., Ohtomo, R., Handa, Y., Kawaguchi, M., Saito, K., et al. (2016). Aquaporin-mediated long-distance polyphosphate translocation directed towards the host in arbuscular mycorrhizal symbiosis: application of virus-induced gene silencing. *New Phytol.* 211, 1202–1208. doi:10.1111/nph.14016.
- Kim, Y. K., Heo, I., and Kim, V. N. (2010). Modifications of Small RNAs and their associated proteins. *Cell* 143, 703–709. doi:10.1016/j.cell.2010.11.018.

- Kloppholz, S., Kuhn, H., and Requena, N. (2011). A secreted fungal effector of *Glomus intraradices* promotes symbiotic biotrophy. *Curr. Biol.* 21, 1204–1209. doi:10.1016/j.cub.2011.06.044.
- Kobae, Y., and Hata, S. (2010). Dynamics of periarbuscular membranes visualized with a fluorescent phosphate transporter in arbuscular mycorrhizal roots of rice. *Plant Cell Physiol.* 51, 341–353. doi:10.1093/pcp/pcq013.
- Kobayashi, Y., Maeda, T., Yamaguchi, K., Kameoka, H., Tanaka, S., Ezawa, T., et al. (2018). The genome of *Rhizophagus clarus* HR1 reveals a common genetic basis for auxotrophy among arbuscular mycorrhizal fungi. *BMC Genomics* 19, 465. doi:10.1186/s12864-018-4853-0.
- Koch, A., Biedenkopf, D., Furch, A., Weber, L., Rossbach, O., Abdellatif, E., et al. (2016). An RNAi-based control of *Fusarium graminearum* infections through spraying of long dsRNAs involves a plant passage and is controlled by the fungal silencing machinery. *PLoS Pathog.* 12, 1–22. doi:10.1371/journal.ppat.1005901.
- Krajinski, F., Courty, P. E., Sieh, D., Franken, P., Zhang, H., Bucher, M., et al. (2014). The H⁺-ATPase HA1 of *Medicago truncatula* is essential for phosphate transport and plant growth during arbuscular mycorrhizal symbiosis. *Plant Cell* 26, 1808–1817. doi:10.1105/tpc.113.120436.
- Kuhn, C. D., and Joshua-Tor, L. (2013). Eukaryotic Argonautes come into focus. *Trends Biochem. Sci.* 38, 263–271. doi:10.1016/j.tibs.2013.02.008.
- Lamonte, G., Philip, N., Reardon, J., Lacsina, J. R., Majoros, W., Chapman, L., et al. (2012). Translocation of sickle cell erythrocyte MicroRNAs into *Plasmodium falciparum* inhibits parasite translation and contributes to malaria resistance. *Cell Host Microbe* 12, 187–199. doi:10.1016/j.chom.2012.06.007.
- Lanfranco, L., Bonfante, P., and Genre, A. (2016). The Mutualistic Interaction between Plants and Arbuscular Mycorrhizal Fungi. *The Fungal Kingdom* 727–747. doi:10.1128/microbiolspec.funk-0012-2016.
- Lanfranco, L., Fiorilli, V., and Gutjahr, C. (2018). Partner communication and role of nutrients in the arbuscular mycorrhizal symbiosis. *New Phytol.* doi:10.1111/nph.15230.
- Lauressergues, D., Delaux, P. M., Formey, D., Lelandais-Brière, C., Fort, S., Cottaz, S., et al. (2012). The microRNA miR171h modulates arbuscular mycorrhizal colonization of *Medicago truncatula* by targeting NSP2. *Plant J.* 72, 512–522. doi:10.1111/j.1365-313X.2012.05099.x.
- Le Marquer, M., Bécard, G., and Frei dit Frey, N. (2019). Arbuscular mycorrhizal fungi possess a CLAVATA3/embryo surrounding region-related gene that positively regulates symbiosis. *New Phytol.* 222, 1030–1042. doi:10.1111/nph.15643.
- Lee, H. C., Chang, S. S., Choudhary, S., Aalto, A. P., Maiti, M., Bamford, D. H., et al. (2009). QiRNA is a new type of small interfering RNA induced by DNA damage. *Nature* 459, 274–277. doi:10.1038/nature08041.
- Lee, H. C., Li, L., Gu, W., Xue, Z., Crosthwaite, S. K., Pertsemliadis, A., et al. (2010). Diverse pathways generate microRNA-like RNAs and dicer-independent small interfering RNAs in fungi. *Mol. Cell* 38, 803–814. doi:10.1016/j.molcel.2010.04.005.

- Lelandais-Brière, C., Moreau, J., Hartmann, C., and Crespi, M. (2016). Noncoding RNAs, emerging regulators in root endosymbioses. *Mol. Plant-Microbe Interact.* 29, 170–180. doi:10.1094/MPMI-10-15-0240-FI.
- Liao, D., Wang, S., Cui, M., Liu, J., Chen, A., and Xu, G. (2018). Phytohormones regulate the development of arbuscular mycorrhizal symbiosis. *Int. J. Mol. Sci.* 19. doi:10.3390/ijms19103146.
- Liu, J., Maldonado-Mendoza, I., Lopez-Meyer, M., Cheung, F., Town, C. D., and Harrison, M. J. (2007). Arbuscular mycorrhizal symbiosis is accompanied by local and systemic alterations in gene expression and an increase in disease resistance in the shoots. *Plant J.* 50, 529–544. doi:10.1111/j.1365-313X.2007.03069.x.
- Liu, S., Da Cunha, A. P., Rezende, R. M., Cialic, R., Wei, Z., Bry, L., et al. (2016). The host shapes the gut microbiota via fecal microRNA. *Cell Host Microbe* 19, 32–43. doi:10.1016/j.chom.2015.12.005.
- Luginbuehl, L. H., Menard, G. N., Kurup, S., Van Erp, H., Radhakrishnan, G. V, Breakspear, A., et al. (2017). Fatty acids in arbuscular mycorrhizal fungi are synthesized by the host plant. *Science.* 356, 1175–1178. doi:10.1126/science.aan0081.
- Luginbuehl, L. H., and Oldroyd, G. E. D. (2017). Understanding the arbuscule at the heart of endomycorrhizal symbioses in plants. *Curr. Biol.* 27, R952–R963. doi:10.1016/j.cub.2017.06.042.
- MacLean, A. M., Bravo, A., and Harrison, M. J. (2017). Plant signaling and metabolic pathways enabling arbuscular mycorrhizal symbiosis. *Plant Cell*, tpc.00555.2017. doi:10.1105/tpc.17.00555.
- MacRae, I. J., and Doudna, J. A. (2007). Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr. Opin. Struct. Biol.* 17, 138–145. doi:10.1016/J.SBI.2006.12.002.
- Manck-Götzenberger, J., and Requena, N. (2016). Arbuscular mycorrhiza symbiosis induces a major transcriptional reprogramming of the potato SWEET sugar transporter family. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.00487.
- Marí-Ordóñez, A., Marchais, A., Etcheverry, M., Martin, A., Colot, V., and Voinnet, O. (2013). Reconstructing de novo silencing of an active plant retrotransposon. *Nat. Genet.* 45, 1029–1039. doi:10.1038/ng.2703.
- Márquez, L. M., Redman, R. S., Rodriguez, R. J., and Roossinck, M. J. (2007). A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. *Science.* 315, 513–515. doi:10.1126/science.1136237.
- Martin, F. M., Uroz, S., and Barker, D. G. (2017). Ancestral alliances: Plant mutualistic symbioses with fungi and bacteria. *Science.* 356. doi:10.1126/science.aad4501.
- Matzke, M. A., and Mosher, R. A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15, 394–408. doi:10.1038/nrg3683.
- Mayoral, J. G., Hussain, M., Albert Joubert, D., Iturbe-Ormaetxe, I., O'Neill, S. L., and Asgari, S. (2014). *Wolbachia* small noncoding RNAs and their role in cross-kingdom communications. *Proc. Natl. Acad. Sci. U. S. A.* 111, 18721–18726. doi:10.1073/pnas.1420131112.

- McLoughlin, A. G., Wytinck, N., Walker, P. L., Girard, I. J., Rashid, K. Y., De Kievit, T., et al. (2018). Identification and application of exogenous dsRNA confers plant protection against *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *Sci. Rep.* 8, 1–14. doi:10.1038/s41598-018-25434-4.
- Milgroom, M. G., and Cortesi, P. (2004). Biological control of chestnut blight with hypovirulence: a critical analysis. *Annu. Rev. Phytopathol.* 42, 311–338. doi:10.1146/annurev.phyto.42.040803.140325.
- Miozzi, L., Vaira, A. M., Catoni, M., Fiorilli, V., Accotto, G. P., and Lanfranco, L. (2019). Arbuscular mycorrhizal symbiosis: plant friend or foe in the fight against viruses? *Front. Microbiol.* 10, 1238. doi:10.3389/fmicb.2019.01238.
- Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. *Nature* 457, 413–420. doi:10.1038/nature07756.
- Morin, E., Miyauchi, S., San Clemente, H., Chen, E. C. H., Pelin, A., Providencia, I., et al. (2019). Comparative genomics of *Rhizophagus irregularis*, *R. cerebriforme*, *R. diaphanus* and *Gigaspora rosea* highlights specific genetic features in Glomeromycotina. *New Phytol.* 222, 1584–1598. doi:10.1111/nph.15687.
- Müller, L. M., and Harrison, M. J. (2019). Phytohormones, miRNAs, and peptide signals integrate plant phosphorus status with arbuscular mycorrhizal symbiosis. *Curr. Opin. Plant Biol.* 50, 132–139. doi:10.1016/j.pbi.2019.05.004.
- Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell* 2, 279–289. doi:10.2307/3869076.
- Nerva, L., Ciuffo, M., Vallino, M., Margaria, P., Varese, G. C., Gnani, G., et al. (2016). Multiple approaches for the detection and characterization of viral and plasmid symbionts from a collection of marine fungi. *Virus Res.* 219, 22–38. doi:10.1016/J.VIRUSRES.2015.10.028.
- Nerva, L., Varese, G. C., and Turina, M. (2018). Different approaches to discover mycovirus associated to marine organisms. *Methods Mol. Biol.* 1746, 97–114. doi:10.1007/978-1-4939-7683-6_8.
- Nicolás, F. E., and Garre, V. (2017). RNA Interference in fungi: retention and loss. *The Fungal Kingdom*, 657–671. doi:10.1128/microbiolspec.funk-0008-2016.
- Nicolás, F. E., Torres-Martínez, S., and Ruiz-Vázquez, R. M. (2013). Loss and retention of RNA interference in fungi and parasites. *PLoS Pathog.* 9, 1–4. doi:10.1371/journal.ppat.1003089.
- Oldroyd, G. E. D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* 11, 252–63. doi:10.1038/nrmicro2990.
- Pandey, P., Wang, M., Baldwin, I. T., Pandey, S. P., and Groten, K. (2018). Complex regulation of microRNAs in roots of competitively-grown isogenic *Nicotiana attenuata* plants with different capacities to interact with arbuscular mycorrhizal fungi. *BMC Genomics* 19, 937. doi:10.1186/s12864-018-5338-x.
- Pawlowska, T. E., Gaspar, M. L., Lastovetsky, O. A., Mondo, S. J., Real-Ramirez, I., Shakya, E., et al. (2018). Biology of fungi and their bacterial endosymbionts. *Annu. Rev. Phytopathol.* 56, 289–309. doi:10.1146/annurev-phyto-080417-045914.

- Pfeffer, P. E., Douds, D. D., Bécard, G., and Shachar-Hill, Y. (1999). Carbon uptake and the metabolism and transport of lipids in an arbuscular mycorrhiza. *Plant Physiol.* 120, 587–598. doi:10.1104/pp.120.2.587.
- Pimprikar, P., and Gutjahr, C. (2018). Transcriptional regulation of arbuscular mycorrhiza development. *Plant Cell Physiol.* 59, 673–690. doi:10.1093/pcp/pcy024.
- Ren, B., Wang, X., Duan, J., and Ma, J. (2019). Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation. *Science.* 365, 919–922. doi:10.1126/science.aav8907.
- Ren, G., Xie, M., Zhang, S., Vinovskis, C., Chen, X., and Yu, B. (2014). Methylation protects microRNAs from an AGO1-associated activity that uridylyates 5' RNA fragments generated by AGO1 cleavage. *Proc. Natl. Acad. Sci. U. S. A.* 111, 6365–6370. doi:10.1073/pnas.1405083111.
- Rich, M. K., Schorderet, M., Bapaume, L., Falquet, L., Morel, P., Vandenbussche, M., et al. (2015). The petunia GRAS transcription factor ATA/RAM1 regulates symbiotic gene expression and fungal morphogenesis in arbuscular mycorrhiza. *Plant Physiol.* 168, 788–797. doi:10.1104/pp.15.00310.
- Romano, N., and Macino, G. (1992). Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Mol. Microbiol.* 6, 3343–3353. doi:10.1111/j.1365-2958.1992.tb02202.x.
- Roossinck, M. J. (2011). The good viruses: viral mutualistic symbioses. *Nat. Rev. Microbiol.* 9, 99–108. doi:10.1038/nrmicro2491.
- Roth, R., and Paszkowski, U. (2017). Plant carbon nourishment of arbuscular mycorrhizal fungi. *Curr. Opin. Plant Biol.* 39, 50–56. doi:10.1016/j.pbi.2017.05.008.
- Roth, R., Hillmer, S., Funaya, C., Chiapello, M., Schumacher, K., Lo Presti, L., et al. (2019). Arbuscular cell invasion coincides with extracellular vesicles and membrane tubules. *Nat. Plants* 5, 204–211. doi:10.1038/s41477-019-0365-4.
- Ruiz-Lozano, J. M. (2003). Arbuscular mycorrhizal symbiosis and alleviation of osmotic stress. New perspectives for molecular studies. *Mycorrhiza* 13, 309–317. doi:10.1007/s00572-003-0237-6.
- Salehin, M., Bagchi, R., and Estelle, M. (2015). SCFTIR1/AFB-based auxin perception: mechanism and role in plant growth and development. *Plant Cell* 27, 9–19. doi:10.1105/tpc.114.133744.
- Schmitz, A. M., and Harrison, M. J. (2014). Signaling events during initiation of arbuscular mycorrhizal symbiosis. *J. Integr. Plant Biol.* 56, 250–261. doi:10.1111/jipb.12155.
- Segers, G. C., Van Wezel, R., Zhang, X., Hong, Y., and Nuss, D. L. (2006). Hypovirus papain-like protease p29 suppresses RNA silencing in the natural fungal host and in a heterologous plant system. *Eukaryot. Cell* 5, 896–904. doi:10.1128/EC.00373-05.
- Segers, G., Zhang, X., Deng, F., Sun, Q., and Nuss, D. L. (2007). Evidence that RNA silencing functions as an antiviral defense mechanism in fungi. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12902–12906. doi:10.1073/pnas.0702500104.
- Shahid, S., Kim, G., Johnson, N. R., Wafula, E., Wang, F., Coruh, C., et al. (2018). MicroRNAs from the parasitic plant *Cuscuta campestris* target host messenger RNAs. *Nature* 553, 82–85. doi:10.1038/nature25027.

- Shiu, P. K. T., Raju, N. B., Zickler, D., and Metzenberg, R. L. (2001). Meiotic silencing by unpaired DNA. *Cell* 107, 905–916. doi:10.1016/S0092-8674(01)00609-2.
- Shiu, P. K. T., Zickler, D., Raju, N. B., Ruprich-Robert, G., and Metzenberg, R. L. (2006). SAD-2 is required for meiotic silencing by unpaired DNA and perinuclear localization of SAD-1 RNA-directed RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* 103, 2243–8. doi:10.1073/pnas.0508896103.
- Smibert, P., Yang, J.-S., Azzam, G., Liu, J.-L., and Lai, E. C. (2013). Homeostatic control of Argonaute stability by microRNA availability. *Nat. Struct. Mol. Biol.* 20, 789–95. doi:10.1038/nsmb.2606.
- Spatafora, J. W., Chang, Y., Benny, G. L., Lazarus, K., Smith, M. E., Berbee, M. L., et al. (2016). A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* 108, 1028–1046. doi:10.3852/16-042.
- Sun, X., Chen, W., Ivanov, S., MacLean, A. M., Wight, H., Ramaraj, T., et al. (2019). Genome and evolution of the arbuscular mycorrhizal fungus *Diversispora epigaea* (formerly *Glomus versiforme*) and its bacterial endosymbionts. *New Phytol.* 221, 1556–1573. doi:10.1111/nph.15472.
- Tang, N., San Clemente, H., Roy, S., Bécard, G., Zhao, B., and Roux, C. (2016). A Survey of the gene repertoire of *Gigaspora rosea* unravels conserved features among glomeromycota for obligate biotrophy. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.00233.
- Teng, Y., Ren, Y., Sayed, M., Hu, X., Lei, C., Kumar, A., et al. (2018). Plant-derived exosomal MicroRNAs shape the gut microbiota. *Cell Host Microbe* 24, 637-652.e8. doi:10.1016/J.CHOM.2018.10.001.
- Torres-Martínez, S., and Ruiz-Vázquez, R. M. (2016). RNAi pathways in *Mucor*: a tale of proteins, small RNAs and functional diversity. *Fungal Genet. Biol.* 90, 44–52. doi:10.1016/j.fgb.2015.11.006.
- Torres-Martínez, S., and Ruiz-Vázquez, R. M. (2017). The RNAi universe in fungi: a varied landscape of small RNAs and biological functions. *Annu. Rev. Microbiol.* 71, 371–391. doi:10.1146/annurev-micro-090816-093352.
- Trieu, T. A., Calo, S., Nicolás, F. E., Vila, A., Moxon, S., Dalmay, T., et al. (2015). A non-canonical RNA silencing pathway promotes mRNA degradation in basal fungi. *PLoS Genet.* 11, 1–32. doi:10.1371/journal.pgen.1005168.
- Turina, M., Ghignone, S., Astolfi, N., Silvestri, A., Bonfante, P., and Lanfranco, L. (2018). The virome of the arbuscular mycorrhizal fungus *Gigaspora margarita* reveals the first report of DNA fragments corresponding to replicating non-retroviral RNA viruses in Fungi. *Environ. Microbiol.* 00. doi:10.1111/1462-2920.14060.
- Venice, F., Ghignone, S., Salvioli, A., Amselem, J., Novero, M., Xianan, X., et al. (2019). At the nexus of three kingdoms: the genome of the mycorrhizal fungus *Gigaspora margarita* provides insights into plant, endobacterial and fungal interactions. 00. doi:10.1111/1462-2920.14827.
- Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I. S., et al. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. *Science.* 303, 672–676. doi:10.1126/science.1093686.

- Vidal, E. A., Araus, V., Lu, C., Parry, G., Green, P. J., Coruzzi, G. M., et al. (2010). Nitrate-responsive miR393/AFB3 regulatory module controls root system architecture in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4477–4482. doi:10.1073/pnas.0909571107.
- Villalobos-Escobedo, J. M., Herrera-Estrella, A., and Carreras-Villaseñor, N. (2016). The interaction of fungi with the environment orchestrated by RNAi. *Mycologia* 108, 556–571. doi:10.3852/15-246.
- Volpe, V., Giovannetti, M., Sun, X.-G., Fiorilli, V., and Bonfante, P. (2016). The phosphate transporters LjPT4 and MtPT4 mediate early root responses to phosphate status in non mycorrhizal roots. *Plant. Cell Environ.* 39, 660–671. doi:10.1111/pce.12659.
- Voß, S., Betz, R., Heidt, S., Corradi, N., and Requena, N. (2018). RiCRN1, a crinkler effector from the arbuscular mycorrhizal fungus *Rhizophagus irregularis*, functions in arbuscule development. *Front. Microbiol.* 9, 1–18. doi:10.3389/fmicb.2018.02068.
- Wang, B., Sun, Y., Song, N., Zhao, M., Liu, R., Feng, H., et al. (2017a). *Puccinia striiformis* f. sp. *tritici* microRNA-like RNA 1 (*Pst-miR1*), an important pathogenicity factor of *Pst*, impairs wheat resistance to *Pst* by suppressing the wheat pathogenesis-related 2 gene. *New Phytol.* 215, 338–350. doi:10.1111/nph.14577.
- Wang, E., Yu, N., Bano, S. A., Liu, C., Miller, A. J., Cousins, D., et al. (2014). A H⁺-ATPase that energizes nutrient uptake during mycorrhizal symbioses in rice and *Medicago truncatula*. *Plant Cell* 26, 1818–1830. doi:10.1105/tpc.113.120527.
- Wang, M., Weiberg, A., Dellota, E., Yamane, D., and Jin, H. (2017b). *Botrytis* small RNA Bc-siR37 suppresses plant defense genes by cross-kingdom RNAi. *RNA Biol.* 14, 421–428. doi:10.1080/15476286.2017.1291112.
- Wang, M., Weiberg, A., Lin, F. M., Thomma, B. P. H. J., Huang, H. Da, and Jin, H. (2016). Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. *Nat. Plants* 2, 16151. doi:10.1038/nplants.2016.151.
- Weiberg, A., Wang, M., Lin, F. M., Zhao, H., Zhang, Z., Kaloshian, I., et al. (2013). Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science*. 342, 118–123. doi:10.1126/science.1239705.
- Wewer, V., Brands, M., and Dörmann, P. (2014). Fatty acid synthesis and lipid metabolism in the obligate biotrophic fungus *Rhizophagus irregularis* during mycorrhization of *Lotus japonicus*. *Plant J.* 79, 398–412. doi:10.1111/tpj.12566.
- Wilson, R. C., and Doudna, J. A. (2013). Molecular mechanisms of RNA interference. *Annu. Rev. Biophys.* 42, 217–239. doi:10.1146/annurev-biophys-083012-130404.
- Wu, P., Wu, Y., Liu, C.-C., Liu, L.-W., Ma, F.-F., Wu, X.-Y., et al. (2016). Identification of Arbuscular Mycorrhiza (AM)-Responsive microRNAs in Tomato. *Front. Plant Sci.* 7, 429. doi:10.3389/fpls.2016.00429.
- Xie, X., Lin, H., Peng, X., Xu, C., Sun, Z., Jiang, K., et al. (2016). Arbuscular mycorrhizal symbiosis requires a phosphate transceptor in the *Gigaspora margarita* fungal symbiont. *Mol. Plant* 9, 1583–1608. doi:10.1016/j.molp.2016.08.011.
- Xue, X.-Y., Zhao, B., Chao, L.-M., Chen, D.-Y., Cui, W.-R., Mao, Y.-B., et al. (2014). Interaction between two timing microRNAs controls trichome distribution in *Arabidopsis*. *PLoS Genet.* 10, e1004266. doi:10.1371/journal.pgen.1004266.

- Yaegashi, H., Yoshikawa, N., Ito, T., and Kanematsu, S. (2013). A mycoreovirus suppresses RNA silencing in the white root rot fungus, *Rosellinia necatrix*. *Virology* 444, 409–416. doi:10.1016/J.VIROL.2013.07.010.
- Zeng, T., Rodriguez-Moreno, L., Mansurkhodzaev, A., Wang, P., van den Berg, W., Gascioli, V., et al. (2019). A LysM effector subverts chitin-triggered immunity to facilitate arbuscular mycorrhizal symbiosis. *New Phytol.* 225, 448–460. doi:10.1111/nph.16245.
- Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human dicer and bacterial RNase III. *Cell* 118, 57–68. doi:10.1016/J.CELL.2004.06.017.
- Zhang, Q., Blaylock, L. A., and Harrison, M. J. (2010). Two *Medicago truncatula* half-ABC transporters are essential for arbuscule development in arbuscular mycorrhizal symbiosis. *Plant Cell* 22, 1483–1497. doi:10.1105/tpc.110.074955.
- Zhang, T., Zhao, Y. L., Zhao, J. H., Wang, S., Jin, Y., Chen, Z. Q., et al. (2016). Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen. *Nat. Plants* 2. doi:10.1038/nplants.2016.153.
- Zouari, I., Salvioli, A., Chialva, M., Novero, M., Miozzi, L., Tenore, G. C., et al. (2014). From root to fruit: RNA-Seq analysis shows that arbuscular mycorrhizal symbiosis may affect tomato fruit metabolism. *BMC Genomics* 15, 221. doi:10.1186/1471-2164-15-221.

CHAPTER 2

2. *In silico* analysis of fungal small RNA accumulation reveals putative plant mRNA targets in the symbiosis between an arbuscular mycorrhizal fungus and its host plant

Alessandro Silvestri¹, Valentina Fiorilli¹, Laura Miozzi², Gian Paolo Accotto², Massimo Turina² and Luisa Lanfranco¹

¹Department of Life Sciences and Systems Biology, University of Torino, Torino, Italy

²Institute for Sustainable Plant Protection, CNR, Torino, Italy.

Published on *BMC Genomics*


Silvestri et al. BMC Genomics (2019)
<https://doi.org/10.1186/s12864-019-5561-0>

RESEARCH ARTICLE

Open Access



In silico analysis of fungal small RNA accumulation reveals putative plant mRNA targets in the symbiosis between an arbuscular mycorrhizal fungus and its host plant

Alessandro Silvestri¹, Valentina Fiorilli¹, Laura Miozzi², Gian Paolo Accotto², Massimo Turina² and Luisa Lanfranco^{1*} 

Abstract

Background: Small RNAs (sRNAs) are short non-coding RNA molecules (20–30 nt) that regulate gene expression at transcriptional or post-transcriptional levels in many eukaryotic organisms, through a mechanism known as RNA interference (RNAi). Recent studies have highlighted that they are also involved in cross-kingdom communication: sRNAs can move across the contact surfaces from “donor” to “receiver” organisms and, once in the host cells of the receiver, they can target specific mRNAs, leading to a modulation of host metabolic pathways and defense responses. Very little is known about RNAi mechanism and sRNAs occurrence in Arbuscular Mycorrhizal Fungi (AMF), an important component of the plant root microbiota that provide several benefits to host plants, such as improved mineral uptake and tolerance to biotic and abiotic stress.

Results: Taking advantage of the available genomic resources for the AMF *Rhizophagus irregularis* we described its putative RNAi machinery, which is characterized by a single *Dicer*-like (*DCL*) gene and an unusual expansion of *Argonaute*-like (*AGO*-like) and *RNA-dependent RNA polymerase* (*RdRp*) gene families. In silico investigations of previously published transcriptomic data and experimental assays carried out in this work provided evidence of gene expression for most of the identified sequences. Focusing on the symbiosis between *R. irregularis* and the model plant *Medicago truncatula*, we characterized the fungal sRNA population, highlighting the occurrence of an active sRNA-generating pathway and the presence of microRNA-like sequences. In silico analyses, supported by host plant degradome data, revealed that several fungal sRNAs have the potential to target *M. truncatula* transcripts, including some specific mRNA already shown to be modulated in roots upon AMF colonization.

Conclusions: The identification of RNAi-related genes, together with the characterization of the sRNAs population, suggest that *R. irregularis* is equipped with a functional sRNA-generating pathway. Moreover, the in silico analysis predicted 237 plant transcripts as putative targets of specific fungal sRNAs suggesting that cross-kingdom post-transcriptional gene silencing may occur during AMF colonization.

Keywords: Arbuscular mycorrhizal fungi, *Rhizophagus*, Small RNAs, microRNA-like, RNA interference

* Correspondence: luisa.lanfranco@unito.it

¹Department of Life Sciences and Systems Biology, University of Torino, Viale P.A. Mattioli 25, 10125 Torino, Italy

Full list of author information is available at the end of the article



Background

Rhizophagus irregularis is a model system for arbuscular mycorrhizal fungi (AMF); it belongs to Glomeromycotina [1], a group of soil fungi able to form a mutualistic symbiosis with the majority of land plants. AMF fungi facilitate the supply of water and nutrients to host plants in return of fixed carbon [2]. However, the beneficial effects of the AM symbiosis go beyond an improved mineral nutrition and includes enhanced tolerance to biotic and abiotic stress [3].

A long history of co-evolution characterizes this unique plant-fungus association where the typical highly branched fungal structures (arbuscules), which develop inside cortical cells, represent a clear sign of the occurrence of fine-tuned regulatory circuits in both partners. Such an intimate colonization of plant tissues relies on an efficient molecular communication system, which occurs before the contact, and on extensive structural and metabolic rearrangements on both plant and fungal sides, which have been only partially described [2, 4]. Transcriptomic studies, mainly focused on plant protein-encoding genes, have been instrumental to describe the molecular reprogramming that the AMF colonization induces in different host plants not only locally (roots; [5–7]) but also systemically (shoot and fruit; [8, 9]) level. Nevertheless, investigations on transcript profiles have been performed to a lower extent also on the AMF [10].

Regulation of gene expression relies on several factors related to transcriptional, post-transcriptional and translational events. The most recently characterized level of regulation relies on the RNA interference (RNAi) mechanism and involves small RNAs (sRNAs): they are short non coding RNA molecules (20–30 nt) that can act at transcriptional or post-transcriptional level in many eukaryotic organisms [11, 12]. Basic enzymatic components of the RNAi response are an RNase III protein, Dicer, that produces sRNAs from double-stranded RNAs (dsRNAs) and an Argonaute (AGO) protein, that uses these sRNAs to guide the selective and sequence-specific degradation, translational inhibition or transcriptional repression of the target [13]. An RNA-dependent RNA polymerase (RdRp) is also used in some organisms (nematodes, fungi and plants) to generate dsRNAs from aberrant RNAs and to amplify the silencing signal [13].

The main function initially ascribed to RNAi was the protection of the genome against transposons and exogenous sequences such as invading viruses or transgenes [12]. Later, it became clear that RNAi is also involved in the production of a variety of endogenous sRNAs, which participate, through the control of gene expression, in the regulation of several endogenous biological functions through the control of gene expression [12].

In the 1990s pioneering studies on the filamentous fungus *Neurospora crassa* were seminal to describe the

phenomenon of RNAi in fungi [14]. Since then, investigations on RNAi components and sRNAs populations have been carried out on other fungi and indicated that many of them possess functional sRNAs while some species, such as *Saccharomyces cerevisiae* and *Ustilago maydis*, lost their RNAi capability [15, 16]. Furthermore, studies on fungal models and plant pathogens have shown that fungi may possess different classes of sRNAs, which are produced by multiple Dicer-dependent and Dicer-independent RNAi pathways [13]. Fungi are thus emerging fascinating systems to study RNAi-related processes and, because of their key position in the eukaryotic tree of life, they could provide insights on the evolution and diversification of RNAi.

Interestingly, recent investigations have highlighted that sRNAs are also involved in cross-kingdom communication [17–25]. In particular, concerning the interactions between plants-fungal pathogens or plants-parasitic plants, sRNAs can move across the contact surface, from “donor” to “receiver” organisms. Once in the host cells, sRNAs can target specific host mRNAs, sometimes triggering secondary sRNA production and thus leading to a modulation of host metabolic pathways and defense responses [26–28]. In case of parasitic/pathogenic organisms these findings are of great interest in light of the development of innovative crop defense strategies [23, 24, 29].

Currently very little is known about AMF RNAi machinery [30] and whether AMF possess a population of functional sRNAs. Furthermore, nothing is known about possible sRNAs trafficking and reciprocal sRNA-mediated communication between AMF and host plants. HIGS (host-induced gene silencing) and VIGS (virus induced gene silencing) have been shown to be successful tools for gene silencing in AMF [31–34] suggesting that RNA movement from the host to the fungus indeed occurs and RNAi-related mechanisms are active in AMF. In addition, it has been recently reported that several plant microRNAs are differentially expressed during the AM symbiosis [35–39]; although their functional roles remain widely unclear, some of them could represent potential candidate mobile sRNAs.

Aim of this work was to characterize the essential components of the RNA-mediated gene silencing machinery in the AMF *R. irregularis*, taking advantage of a newly published genome assembly [40]), and to characterize the population of *R. irregularis* sRNAs from extraradical mycelium and symbiotic tissues. We demonstrated that *R. irregularis* possesses key components of the RNAi machinery characterized by an unusual expansion of AGO-like (*Argonaute*-like) and *RdRp* gene families; furthermore, AMF sRNAs share structural properties with previously analyzed fungal sRNA datasets, including microRNA-like sequences. Finally, we identified in silico a list of predicted fungal sRNA-plant host mRNA target pairs possibly involved in

cross-kingdom post-transcriptional gene silencing (PTGS) regulation during AMF colonization.

Results

RNAi machinery in *R. irregularis*

A survey of recently published genomic resources of the AMF *R. irregularis* [40] was performed to identify proteins belonging to the core eukaryotic RNAi machinery: Dicer-like (DCL), AGO and RdRp [13]. By keywords searches on JGI MycCosm portal [41], we found 1 DCL, 40 AGO-like and 21 RdRp putative homologous proteins that responded to the following criteria: the presence of two RNase III domains for DCL [42], the presence of a piwi domain for AGO-like proteins, the presence of an RdRp domain for the RdRp [43]. A blastp search on the predicted *R. irregularis* proteome, using characterized DCL, AGO and RdRp from other fungi (the closely related *Mucor circinelloides* and the RNAi model systems *Neurospora crassa* and *Cryphonectria parasitica*) as queries, resulted in the same number of sequences obtained by keywords searches.

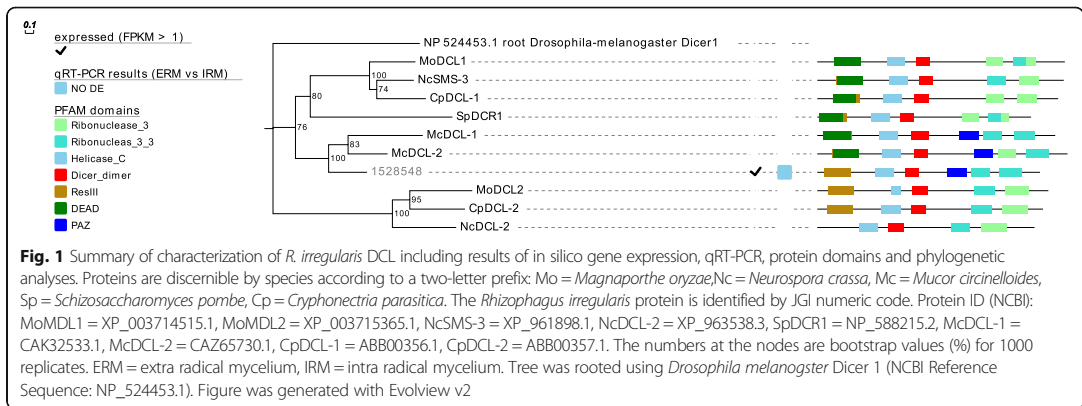
To further characterize the identified sequences, phylogenetic analyses were carried out. A first analysis, performed on DCL proteins, revealed that the only DCL of *R. irregularis* (1528548) is closely related to the two DCL described in *M. circinelloides* [44], consistent with the evolutionary relationships of the two taxonomic groups [1] (Fig. 1) and confirming the analysis carried out by Lee et al. (2018) on a previous, more fragmented, version of *R. irregularis* genome assembly [30]. Interestingly, Lee et al. (2018) also identified two additional prokaryotic (class I) ribonuclease III protein coding genes, which seem to derive from horizontal gene transfer from cyanobacteria [30].

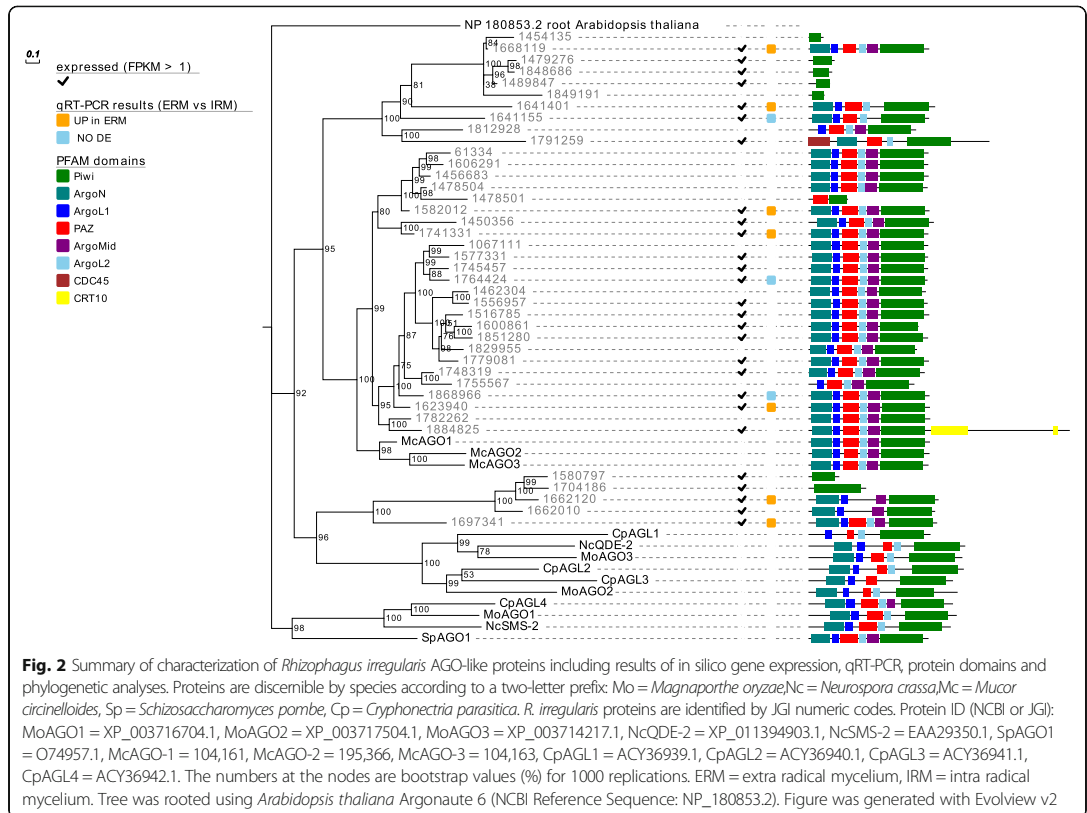
Regarding AGO, 25 of the 40 AGO-like sequences, possessed all the 4 typical AGO core domains - piwi, PAZ, MID and N-terminal [45] - whereas the remaining

15 lacked some of the non-piwi domains present in typical AGO (Fig. 2). A phylogenetic analysis of the identified AGO-like sequences revealed that the *R. irregularis* genome encodes for 5 proteins (1580797, 1704186, 1662120, 1662010, 1697341) related to AGO of fungi belonging to Ascomycota (*M. oryzae*, *N. crassa*, *C. parasitica* and *S. pombe*), while 25 proteins (61334, 1606291, 1456683, 1478504, 1478501, 1582012, 1450356, 1741331, 1067111, 1577331, 1745457, 1764424, 1462304, 1556957, 1516785, 1600861, 1851280, 1829955, 1779081, 1748319, 1755567, 1868966, 1623940, 1782262 and 1884824) form a group with the three AGO proteins from *M. circinelloides*, a fungus which belongs to the Mucoromycota phylum that also includes AMF [1], and for which the RNAi machinery has been well characterized [13, 44].

The AGO gene family is divided in three paralogous groups: a widespread AGO-like group found in plants, animals and fungi, a Piwi-like group closely related to *Drosophila melanogaster* PIWI (P-element Induced Wimpy Testis) only found in animals, and a species-specific group (group 3 AGO) only found in *Caenorhabditis elegans* [46]. Interestingly, the genome of *C. elegans* also displays the highest level of AGO gene expansion so far reported (26 total genes; [46]). Considering that a similar degree of expansion is also observed in *R. irregularis*, we wondered if some of the identified *R. irregularis* AGO-like proteins were related with those of the animal Piwi-like group or with the ones specific of the *C. elegans* group 3. For this purpose, a phylogenetic tree derived including *D. melanogaster*, *C. elegans* and *Arabidopsis thaliana* AGO did not reveal any homologous of group 3 AGO (those specific of *C. elegans*) or of Piwi-like AGO in *R. irregularis* (Additional file 1: Figure S1).

In addition, our bioinformatics search allowed the identification of 21 putative RdRp proteins. The phylogenetic analysis shows that well characterized RdRp from Ascomycetes are grouped in three clades (Fig. 3).





Fifteen *R. irregularis* proteins cluster within the clade containing RdRp1 from *Magnaporthe oryzae* and are more related to the two proteins from *M. circinelloides* (144762, 135684). This suggests that these 15 sequences may be a product of a recent gene expansion event. Three *R. irregularis* sequences (1778075, 1581910, 1697445) are grouped together with the clade containing RdRp2 from *M. oryzae*, close to the *M. circinelloides* 82,874 sequence. The association of the last three RdRp proteins (1473733, 1669713, 1646639) to the clade containing *M. oryzae* RdRp3 is not statistically well supported (Fig. 3). When we added plant RdRp from the model organism *A. thaliana* in the analysis, no difference in the structure of the tree topology was detected (Additional file 1: Figure S2).

We wondered if a similar occurrence of RNAi-related genes is noticeable in other AMF. In the recently published *Rhizophagus clarus* proteome [47] we found 2 putative DCL, 33 putative AGO-like and 17 putative RdRp. Phylogenetic analyses of *R. clarus* AGO-like, RdRp and DCL proteins revealed that they are strictly related with those of *R. irregularis* (Additional file 1: Figure S3).

To find evidence of gene expression of the putative RNAi machinery, publicly available RNA-seq data [40] obtained from *R. irregularis* germinating spores and symbiotic tissues (mycorrhizal roots) were analyzed. Interestingly, the *DCL* gene (Fig. 1), 27 out of the 40 AGO-like genes (Fig. 2) and 19 out of 21 *RdRp* genes (Fig. 3) are all expressed in at least one of two considered conditions (Additional file 2). To support the in silico expression analyses, we performed quantitative RT-PCR (qRT-PCR) assays on 13 genes (the single *DCL*, 10 AGO-like and 2 *RdRp*), randomly chosen from the group of expressed sequences. We focused on the symbiotic phase of the *M. truncatula*-*R. irregularis* association, considering the extraradical mycelium (ERM) and the intraradical mycelium (IRM), obtained by removing under a stereomicroscope the ERM from mycorrhizal roots. Seven AGO-like and 1 *RdRp* mRNAs (1662120, 1697341, 1582012, 1741331, 1623940, 1641401, 1668119 and 1778075) were up-regulated in ERM compared to IRM, while 3 AGO-like, 1 *RdRp* and the *DCL* (1868966, 1764424, 1641155, 1578121 and 1528548) showed no differential expression in the two conditions tested (Fig. 4).

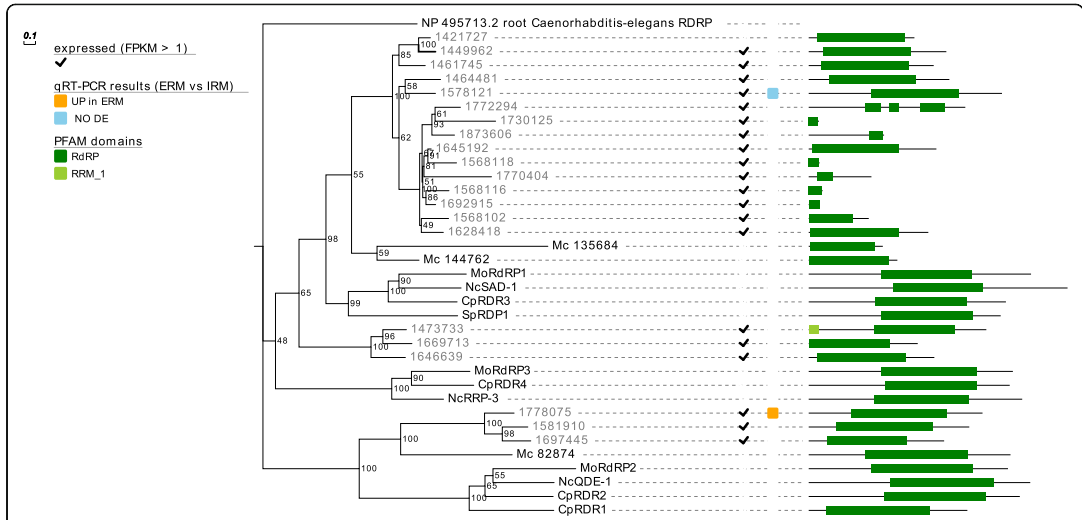


Fig. 3 Summary of characterization of *Rhizophagus irregularis* RdRp including results of in silico gene expression, qRT-PCR, protein domains and phylogenetic analyses. Proteins are discernible by species according to a two-letter prefix Mo = *Magnaporthe oryzae*, Nc = *Neurospora crassa*, Mc = *Mucor circinelloides*, Sp = *Schizosaccharomyces pombe*, Cp = *Cryphonectria parasitica*. *R. irregularis* proteins are identified by JGI numeric codes. Protein ID (NCBI or JGI): MoRdRP1 = XP_003721007.1, MoRdRP2 = XP_003711624.1, MoRdRP3 = XP_003712093.1, NcQDE-1 = EAA29811.1, NcSAD-1 = XP_964248.3, NcRRP-3 = XP_963405.1, SpRDP1 = NP_001342838.1, McRdRP-1 = 111,871, McRdRP-2 = 104,159, CpRDR1 = 270,014, CpRDR2 = 35,624, CpRDR3 = 10,929, CpRDR4 = 339,656. The numbers at the nodes are bootstrap values (%) for 1000 replications. ERM = extra radical mycelium, IRM = intra radical mycelium. Tree was rooted using *Caenorhabditis elegans* RDRP (NCBI Reference Sequence: NP_495713.2). Figure was generated with Evolview v2

Characterization of small RNAs

To characterize the *R. irregularis* sRNA population, we sequenced, with an Illumina platform, 9 sRNAs libraries prepared from biological samples in different conditions of the *R. irregularis* - *M. truncatula* symbiotic association: 3 from extraradical mycelium (ERM; fungal

structures developing outside the roots after colonization), 3 from mycorrhizal roots from which we removed the extraradical mycelium (RM) and 3 from non mycorrhizal roots (RC). The presence of a functional AM symbiosis in RM samples was confirmed by qRT-PCR assays using primers for the plant

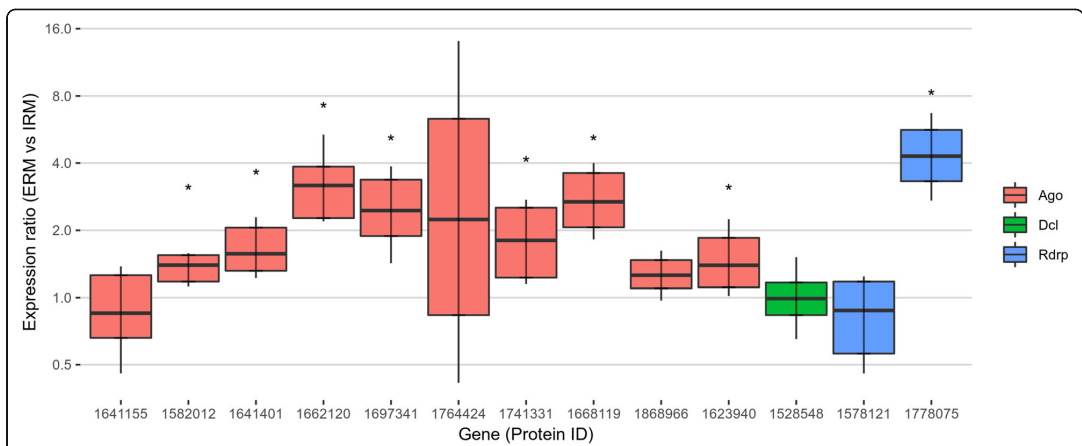


Fig. 4 Whisker-box plot of the relative expression (ERM = “extraradical mycelium” vs IRM = “intraradical mycelium”) calculated by REST2009 software of 10 AGO-like, 2 *RdRp* and the *DCL* genes identified in *Rhizophagus irregularis*, here reported with their JGI protein ID. Asterisks highlight genes with significant differential expression between the two conditions

AMF-inducible phosphate transporter gene (*MtPT4*) (Additional file 1: Figure S4).

A total of 229,660,397 reads were generated; after adapter removal and filtering for quality, artifacts, tRNA, rRNA, snRNA and snoRNA presence, 53,746,056 were retained (Additional file 3). Reads were then mapped on *M. truncatula* and *R. irregularis* genomes allowing zero mismatches. Considering the different biological replicates, the 76–82% of reads from ERM libraries mapped on the fungal genome and less than 1% on the plant genome, probably because of a contamination by root material during ERM harvesting (Fig. 5a), even though we can not exclude a possible plant-originated sRNA component present natively in ERM as recently observed in the *Botrytis cinerea*-host plant interaction [27]. The 76–85% of reads from RC libraries mapped on plant genome with a very limited number of reads mapping on fungal genome (0.01–0.02%). For RM samples an intermediate situation was observed with 62–70% reads mapping on plant genome and 10–20% on fungal genome. A very low percentage of reads for each condition mapped on both plant and fungal genomes: about 0.1% for ERM, 0.01% for RC and 0.02–0.05% for RM libraries (Fig. 5a).

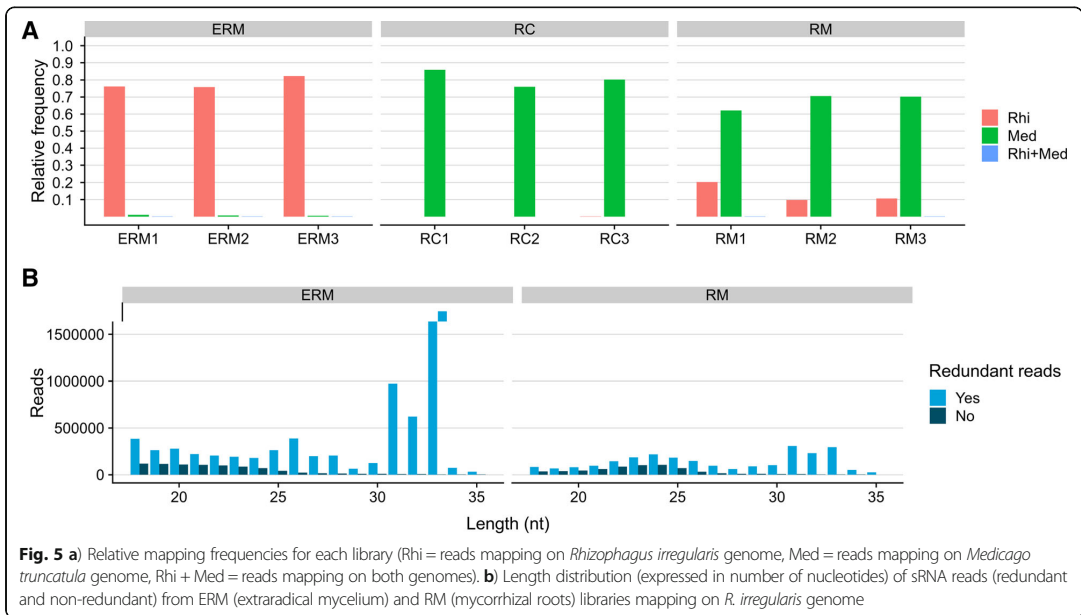
The evaluation of read length distribution is a useful tool to assess whether sRNAs are originated through a specific molecular pathway [48], i.e. in a Dicer-dependent manner. Plant reads present in RM and RC libraries displayed a typical enrichment of 21 and 24 nt-long sequences [37, 49], with the 21 nt-long class consisting of more redundant sequences than the 24 nt-long class

(Additional file 1: Figure S5). On the contrary, the length distribution of sRNA reads from ERM and RM libraries mapping on the fungal genome (*R. irregularis* sRNAs = *Rir*-sRNAs) was bimodal with a first peak at 24 nt and 26 nt in RM and ERM respectively, and a second peak at 31–32-33 nt in both samples; the 31–33 nt long reads consist of extremely redundant sequences (Fig. 5b).

The analysis of the 5' terminal nucleotide composition of fungal non redundant reads showed that approximately half of the sRNAs shorter than 26 nt starts with uracil (Additional file 1: Figure S6). Interestingly, in plants, 5' U enrichment has been associated to the selective loading of sRNAs to specific AGO proteins [50]. The features of these sRNAs, together with the identification of RNAi-related genes, suggest the presence of an active sRNAs-generating pathway in *R. irregularis*.

Characterization of *R. irregularis* sRNA-generating loci

Rir-sRNAs from RM and ERM libraries were used for a genome-guided sRNA-generating loci discovery and characterization, by ShortStack software [51]. Setting a cut-off of 10 RPM (reads per million reads), 2131 sRNA-generating loci, defined by the 95% of *Rir*-sRNAs, were predicted (whole characterization data in Additional file 4). Thirty three percent (702) of *Rir*-sRNA-generating loci localized in intergenic regions while the remaining 67% (1429) shared, for at least one nucleotide, the same genomic coordinates of annotated genes (protein-coding genes). We observed that 69% of *Rir*-sRNAs-generating loci overlapping with annotated



genes produced sRNAs from the same strand of the overlapped genes, 8% from the opposite strand while 7%, despite being located on a specific genomic strand, were localized in regions coding for genes on both strands (so we could not assess if they are sense or anti-sense to genes). The remaining 16% of loci produced sRNAs on both genomic strands. These observations are in line with the results obtained for *M. circinelloides*, where exons were the major source of sRNAs [43]. Differential expression analysis, performed with DESeq2 [52], revealed that 225 *Rir*-sRNAs-generating loci were up-regulated in terms of sRNAs production in ERM while 589 of them were up-regulated in RM; the remaining 1317 loci were not regulated between the two conditions (Additional file 1: Figure S7A).

Considering that transposable elements are an important source for sRNA production [12], we looked for similarity of *Rir*-sRNA-generating loci with fungal repetitive elements from RepBase 23.04 [53]. A total of 236 loci, representing the 11% of the identified loci, had strong similarity with transposons: 93 with DNA transposons, 61 with LTR retrotransposons and 22 with non-LTR retrotransposons. *Rir*-sRNAs mapping on these loci were enriched in 24 nt long sequences (Additional file 1: Figure S8).

In order to evaluate the possible existence of different populations of *Rir*-sRNA-generating loci, we performed a Principal Component Analysis (PCA). For this purpose, as proposed by Fahlgren et al. (2013) [54], we used a model considering the following independent variables: the length of loci, the total number of mapped reads and the nucleotide size proportion of *Rir*-sRNAs (from 18 nt to 35 nt) defining each locus (19 variables in total). The first principal component (PC1), that explained the 25.8% of the total variance, mainly based on the proportion of 21–25 and 27–35 nt-long reads, differentiated the *Rir*-sRNA-generating loci in 2 groups (Fig. 6a). The use of DBSCAN (density-based spatial clustering of applications with noise) algorithm [55] confirmed indeed the presence of two different groups of data: cluster 1 and cluster 2 composed of 1100 and 819 loci, respectively (Fig. 6b). The average nucleotide size distribution of the reads for loci belonging to cluster 1 revealed a decreasing curve from 18 to 35 nt with no evident peaks while for cluster 2 we recorded an enrichment in 22–24 nt-long sequences (Fig. 7).

Interestingly, the two clusters differentiated also on the basis of the genomic positions of the *Rir*-sRNA-generating loci relative to protein-encoding annotated genes (Fig. 6c) and on the basis of the expression levels of the loci between ERM and RM conditions (Fig. 6d). In fact, 94% of *Rir*-sRNA-generating loci from cluster 1 localized in genic regions and 15 and 1% were up-regulated in ERM and RM, respectively, while 37% of loci from cluster 2 localized in genic regions and 1 and 66% were up-regulated in ERM and RM, respectively.

Finally, we observed that 4 and 16% of loci from cluster 1 and cluster 2 respectively showed homology with sequences in RepBase (21 DNA transposons, 12 LTR retrotransposons and 7 non-LTR retrotransposons for cluster 1; 69 DNA transposons, 46 LTR retrotransposons and 15 non-LTR retrotransposons for cluster 2).

***R. irregularis* generates putative miRNA-like sequences**

The ShortStack software predicted 10 *Rir*-sRNA-generating loci as miRNA-like (loci 338, 339, 340, 341, 342, 343, 345, 818, 828 and 1596; Table 1). These sequences, if transcribed, have the ability to form hairpin structures and the software predicts the accumulation of miRNA-miRNA* pairs (Fig. 8). The length of these loci varies from 102 nt (locus 338) to 610 nt (locus 828) while their expression (considering the sum of all RM and ERM libraries) ranges from 196 reads (locus 1596) to 74,526 reads (locus 340; the fourth most expressed sRNA-generating locus). Interestingly, 7 loci are located on the negative strand of the same genome scaffold (scaffold 28) in sequential order within a 8.3 kbp region. The length distribution of reads produced by the *Rir*-miRNA-like loci, as well as the length of mature miRNA sequences, are enriched in sequences from 19 nt to 24 nt (data not shown). Three *Rir*-miRNA-like loci (341, 342, 828) show an increased sRNA production in RM condition.

Identification of *Rir*-sRNAs potentially targeting *M.*

***truncatula* transcripts**

Considering that cross-kingdom RNA silencing seems to be a quite common and widespread phenomenon [17, 21, 25], we searched for in silico evidences of *M. truncatula* mRNAs (*Mtr*-mRNAs) potentially targeted by *Rir*-sRNAs. We combined our sRNA-seq results with degradome (PARE-seq) data collected in similar experimental conditions for the *M. truncatula*-*R. irregularis* symbiotic association (mycorrhizal and non-mycorrhizal roots; [37]). For this purpose, we pooled together all the *Rir*-sRNAs from ERM and RM libraries and counted, for each individual sequence, the number of total occurrences. The 11,396 most abundant *Rir*-sRNAs (expression ranges from 1,945,411 reads for the most abundant *Rir*-sRNA to 19 reads for the less abundant) were used for target prediction against *M. truncatula* transcriptome followed by PARE validation using sPARTA [56].

Resulting targets were further filtered maintaining only the predictions that met the three following criteria: *Rir*-sRNA size between 21 and 24 nt (considering that plant sRNAs involved in RNAi are 21–24 nt long; [57]), adjusted *p*-value less than 0.05, at least 5 PARE reads at cleavage site in mycorrhizal condition and no reads at the same site in non-mycorrhizal one (to limit the search to cleavage signals specific for mycorrhizal condition).

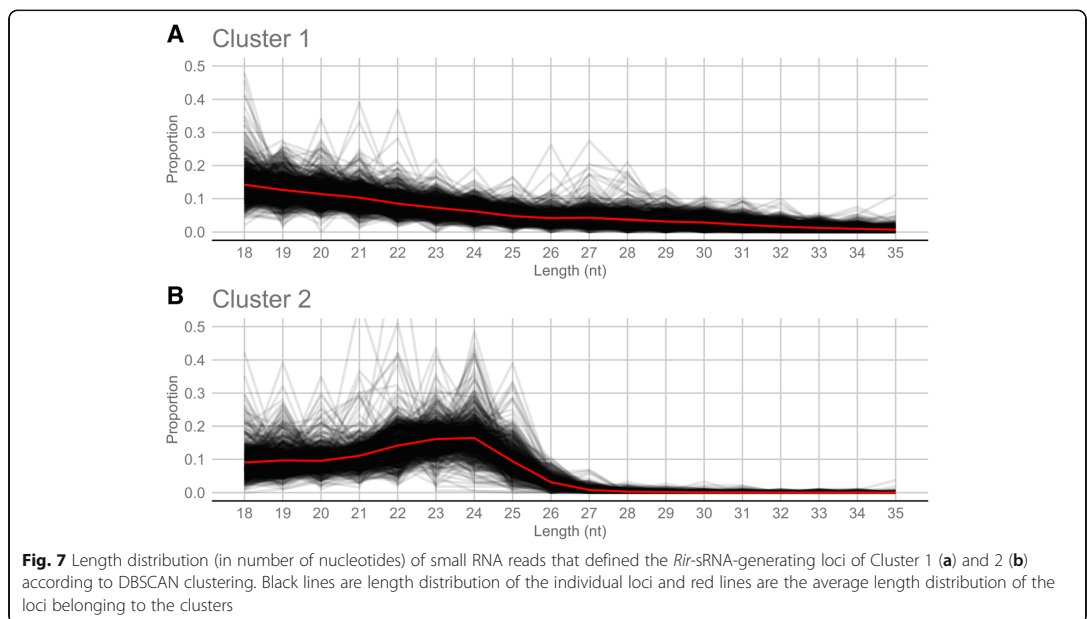
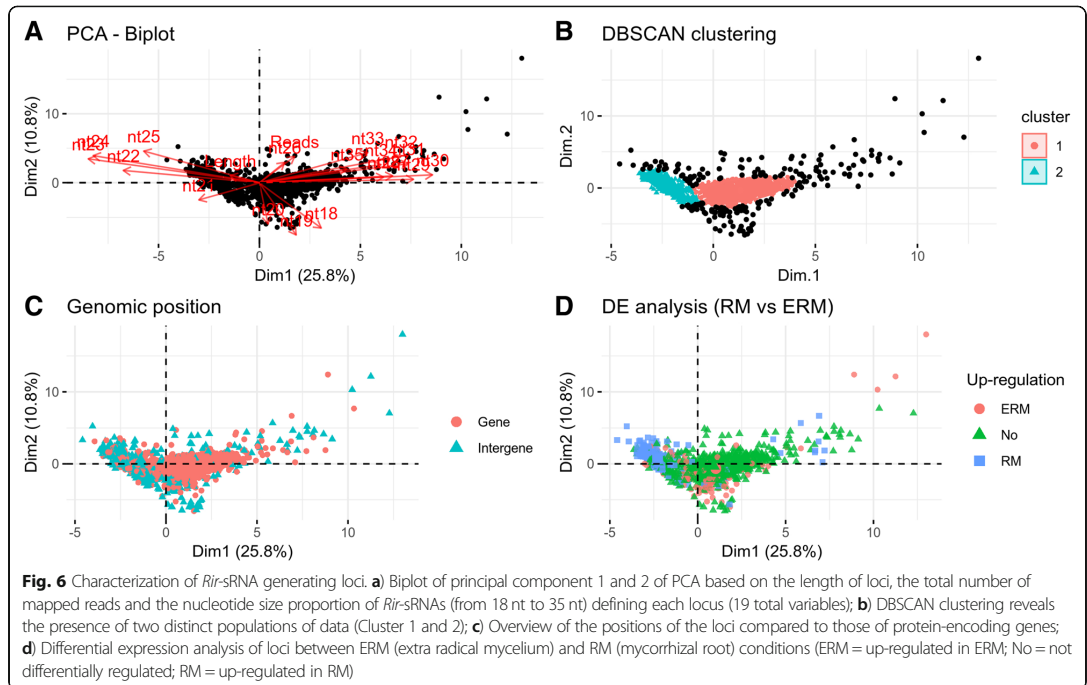


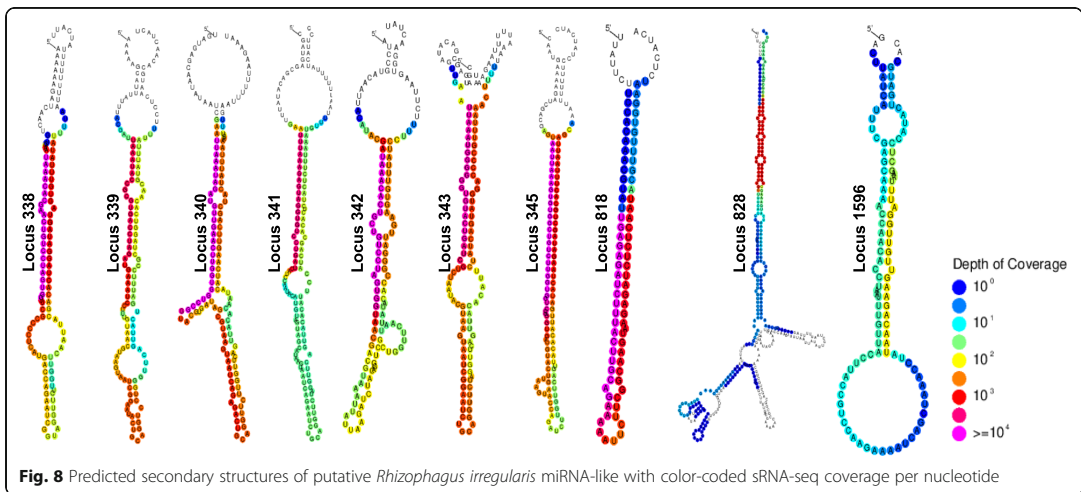
Table 1 Characteristics of predicted miRNA-like loci

Genomic position	Name	Length	Reads	Strand	Mature miRNA-like	Mature miRNA-like length (nt)	Up-regulation
scaffold_28:345427–345,528	Locus_338	102	37,395	–	UAAACACGAACUGCCUAGU	20	No
scaffold_28:346121–346,253	Locus_339	133	2778	–	UAAAUACCGCGUGACCUAGA	20	No
scaffold_28:347535–347,650	Locus_340	116	74,526	–	UUAAAUAGAUGUUGAACUUGGUG	23	No
scaffold_28:349765–349,990	Locus_341	226	3636	–	UUUAAAAGAGUAGGUGUCCUGAUC	23	RM
scaffold_28:350996–351,184	Locus_342	189	9875	–	UAAACACUGCUGCCUAGUGG	21	RM
scaffold_28:351831–351,938	Locus_343	108	7174	–	UUAAAUGGGGGGUGUACUG	19	No
scaffold_28:353658–353,769	Locus_345	112	62,117	–	AAUUAAGUGGGCUGUCUUGGUG	24	No
scaffold_81:287993–288,405	Locus_818	413	17,160	+	UGAGAGAUUUUACUUGCAG	20	No
scaffold_81:370577–371,186	Locus_828	610	2143	–	CGAGGAUCGAGAGCUUGCACGUCA	24	RM
scaffold_245:162243–162,808	Locus_1596	566	196	–	UACAGAAGUUGUUGGAUU	19	No

This analysis identified 310 *Rir*-sRNA - *Mtr*-mRNA interactions involving 237 *Mtr*-mRNA and 274 *Rir*-sRNA (Additional file 5). These *Rir*-sRNAs were mainly produced by *Rir*-sRNA-generating loci up-regulated in RM condition (120), while the remaining corresponded to loci up-regulated in ERM condition (19) or not regulated (47) (Additional file 1: Figure S7B); 4 *Rir*-sRNAs originated from genome regions not annotated as *Rir*-sRNA-generating loci (Additional file 5). Putative targeted *Mtr*-mRNAs were enriched in 8 gene ontology (GO) terms: serine hydrolase activity (GO:0017171) and serine-type peptidase activity (GO:0008236) for the “molecular function” ontology, while cytoplasm (GO:0005737), Golgi apparatus (GO:0005794), Golgi apparatus part (GO:0044431), Golgi membrane (GO:0000139), cytoplasmic part (GO:0044444) and endomembrane system (GO:0012505) for the “cellular components” ontology.

To understand if the *Mtr*-mRNAs potentially targeted by *Rir*-sRNAs can also be targeted by endogenous plant sRNAs, we performed a target prediction and PARE validation against *M. truncatula* transcriptome using as queries the *M. truncatula* miRNAs (*Mtr*-miRNAs) (miRBase, Release 22, [58]). Considering only the predictions with adjusted p-value less than 0.05 and at least 5 PARE reads at cleavage sites in mycorrhizal condition, we identified 296 *Mtr*-miRNA / *Mtr*-mRNA interactions involving 172 *Mtr*-miRNAs and 165 *Mtr*-mRNAs (Additional file 6). According to our criteria, eleven *Mtr*-mRNAs were potentially targeted by both *Rir*-sRNAs and *Mtr*-miRNAs (AES68798, AES68809, AES88206, AES68814, AES74320, AES92729, AES98787, AET00614, KEH18078, KEH21177 and KEH27629) (Additional file 7).

Plant miRNAs (generally the 22 nt-long ones) can trigger secondary phased small interfering RNAs (siRNA) production from their target transcripts [59]. Such a



phenomenon was also observed for some *Cuscuta campestris* miRNAs involved in host-gene regulation by cross-kingdom RNA silencing [28]. To understand whether *Rir*-sRNAs can trigger secondary siRNA production from their in silico predicted *Mtr*-mRNA targets, we followed the approach proposed by Sahid et al. (2018) [28]. After mapping sRNA reads from RM and RC libraries on *M. truncatula* transcriptome, we performed a differential expression analysis that resulted in 575 *Mtr*-mRNAs with an increased number of mapped sRNAs from RM libraries compared to RC ones (Additional file 8). Seven out of these 575 transcripts belong to the group of 237 *Mtr*-mRNAs previously identified as putative targets for *Rir*-sRNAs (AES67976, AES71586, AES75437, AES94149, AET03346, KEH31350 and KEH43815). According to PhaseTank, none of them produced phased siRNAs [60].

Discussion

R. irregularis is equipped with a putative RNAi machinery characterized by the expansion of AGO-like and RdRp

AMF are nowadays recognized as a crucial component of the beneficial plant microbiota. Although their nature of obligate biotrophs has been an important obstacle for their molecular characterization, recent advances in 'omics techniques have allowed to obtain important insights on their biology and evolution [61]. The availability of genome sequences and transcriptomic data covering different fungal life stages gave us the opportunity to study a still unexplored aspect of AMF, that is the prediction of the existence of the RNAi machinery, a key platform for endogenous gene regulation and a possible source of cross-kingdom RNA silencing [10, 40, 47, 62, 63].

By means of blastp- and keywords-based searches, we identified 1 DCL, 40 AGO-like and 21 putative RdRp protein homologues in the genomic resources of the AMF *R. irregularis* [40]. Fungi typically possess only 1–2 DCL, 1–4 AGO and 1–4 RdRp [64], therefore, this high number of AGO-like and RdRp coding genes is unusual. The comparison with the phylogenetically closely related *R. clarus* helped in pointing out the level of conservation among the different AGO, RdRp and DCL proteins.

Interestingly, beside the canonical DCL protein Lee et al. (2018) identified in *R. irregularis* two additional prokaryotic (class I) ribonuclease III proteins (RIRNC2 and RIRNC3) that may arise by putative horizontal gene transfer events from cyanobacteria [30]. It would be interesting to understand whether these two proteins are also functional in the processing of dsRNA.

Concerning AGO, we identified 7 AGO-like proteins that consist of small proteins (peptides) containing only the piwi domain, and that therefore are likely not functional as the AGO protein generally involved in sRNA processing in fungi. Nevertheless, at least the genes

encoding for 4 of them (1479276, 1848686, 1489847 and 1580797) are expressed at sufficiently high levels (Additional file 2) to hypothesize that they are indeed functional, possibly belonging to a new class of non-AGO but piwi domain-containing small peptides, with no evident conserved correspondence into the predicted *R. clarus* proteome. Since functional fungal AGO involved in RNAi have at least 5 or 6 domains (see Fig. 2), we can conclude that *R. irregularis* may possess 25 complete AGO. Among them, 5 (61334, 1606291, 1456683, 1478504, 1478501) do not have specific homologs in the *R. clarus* genome, thus representing a species-specific subclade; however, at the moment there is no evidence of corresponding expressed sequences in transcriptomic databases (Additional file 2).

Comparing both *Rhizophagus* species and *M. circinelloides*, our AGO phylogenetic analysis reveals another interesting aspect: both *Rhizophagus* species have a group of 5–6 AGO-like proteins that are in a clade with well-characterized AGO from ascomycetes, whereas *M. circinelloides* seem to have lost such AGO. The three *M. circinelloides* AGO are in a well-supported clade with 25 AGO-like proteins from *Rhizophagus* displaying a specific AMF expansion of such AGO-like clade. Wider comparisons that include AGO from other eukaryotic kingdoms (plants and animals in specific) help to exclude that any of the fungal AGO are homologues of the PIWI-AGO present in insects (piwi and aubergine from *Drosophila melanogaster*) and point to a specific expansion of AMF specific AGO-like protein clade, well separated from plant and animal AGO.

Regarding RdRp, the discrepancy in the number of RdRp between our work and Lee et al. (2018) -21 vs 3- is likely due to the different criteria used to identify RdRp: while we looked for proteins showing the presence of an RdRp domain, Lee et al. (2018) used rather restrictive parameters to find homologs of RdRp proteins characterized from other fungi (but excluding other RdRp more specific of AMF fungi) [30]. Indeed, following our approach, a similar number of RdRp (17) were also found in the *R. clarus* predicted proteome confirming the conservation of this group of AMF-specific and RdRp motif-containing proteins.

In analogy to what happens with AGO-like, at least 6, out of the 21 identified, are constituted of small peptides that have only the conserved RdRp domains (1730125, 1645192, 1568118, 1770404, 1568116, and 1692915): in this case, all 6 are consistently expressed in transcriptome datasets, and one of them has a specific homologue in *R. clarus*. Also in this case, we can hypothesize the existence of a number of functional RdRp domain-carrying small proteins, so far undescribed in other fungi.

In the case of the phylogenetic tree that includes plant RdRp, it is apparent that orthologues of AMF, fungal

and plants are present, and only a subset of RdRp are specific for AMF (Additional file 1: Figure S2).

Overall, the expansion of AMF-specific AGO-like and RdRp protein gene families prevents us from drawing any conclusion based on functional homology to animal, fungal or plant proteins containing the same conserved domains. In this respect, we can hypothesize that some of these proteins are not involved in RNAi or RNA silencing pathways, but might have a completely new set of functions even unrelated to RNA processing. Expansion of gene families is often accompanied by functional differentiation and gene expression fine-tuning. In this respect, it is tempting to correlate AGO-like and RdRp expansion in AMF to their large genomes, rich in repetitive DNA, mainly transposable elements. This expansion could be the result of a co-evolution between specific anti-transposable element defense and a diversity of transposons present in the genome. It is worth to mention that an evolutionary related AGO enzyme in prokaryotes was shown to protect its host against mobile genetic elements through DNA-guided DNA interference [65]. The setup of a reverse genetic system (i.e. HIGS and VIGS) based on delivery of dsRNA or sRNA for AMF genes would possibly help understanding some of the functional specificities associated to each of the AGO-like or RdRp present in the genome.

In silico transcriptome analyses provided evidence of gene expression for the majority (*DCL*, 27 out of the 40 AGO-like and 19 out of 21 *RdRp* sequences) of the identified RNAi-related genes. Moreover, targeted gene expression profiles showed that some AGO-like and *RdRp* genes are differentially expressed between the extraradical and the intraradical mycelium two functionally distinct compartments of the symbiotic phase at the interface with the soil and with the plant, respectively, supporting the interesting possibility of a functional differentiation among distinct AGO-like and *RdRp* genes in the control of the symbiotic process.

***R. irregularis* sRNAome is characterized by 2 different populations of *Rir*-sRNA-generating loci and by the existence of miRNA-like sequences**

Small RNAseq data generated for the three biological conditions (ERM, RM and RC) were of good quality with relatively high genome-mapping percentages. As expected, enrichment of 21 and 24 nt-long sequences was observed in plant reads of RM and RC samples [37, 49] which proves the good quality of the sRNA and corresponding libraries preparations. Here the attention focused to the fungal sRNAome as no data are currently available for AMF. The length distribution of *R. irregularis* sRNAs (*Rir*-sRNAs) with two distinct peaks clearly differs from a flat curve over 20 nt typically observed in organisms not provided with RNAi machinery [16] or fungal *dcl* knock-out mutants [66].

The characterization of *Rir*-sRNA-generating loci suggests that *R. irregularis* possesses at least two different populations of sRNA-generating loci: the first one (cluster 1) mostly includes sequences overlapping with protein-coding genes, mainly not differentially expressed between ERM and RM, and that produce sRNAs of different sizes, while the second one (cluster 2) is enriched in 22–24 nt-long sRNAs from intergenic sequences, mainly up-regulated in RM. Based on these results, we hypothesize that at least two distinct molecular pathways could contribute to the production of *Rir*-sRNAs. Indeed, two pathways are involved in sRNA generation in *M. circinelloides*: a Dicer-dependent and a Dicer-independent but RdRp-dependent one. Moreover, in analogy to what we observed, protein-encoding genes are the major source for sRNAs production in *M. circinelloides* [13, 44, 67]. Since in both *R. irregularis* and *M. circinelloides* the protein-encoding genes are the major source for sRNA, we could speculate that the *Rir*-sRNA originated from protein-coding sequences are involved in post-transcriptional regulation of the gene from which they originate, as it has been demonstrated in *M. circinelloides* [68, 69]. We did not perform a *Rir*-sRNA target prediction on fungal endogenous genes since no software for sRNAs target prediction in fungi is available and the predictions made with software developed for other organisms have rarely been experimentally validated [13].

Among the population of *Rir*-sRNA-generating loci, 10 were predicted as miRNA-like. So far, miRNA-like have been identified in several fungi belonging to Ascomycota and Basidiomycota phyla but not in the basal fungus *M. circinelloides* [13, 69, 70]. This is the first evidence of miRNA-like occurrence in the Mucoromycota group. Since no miRNA-like database is currently available for fungal sequences, analysis for homolog sequences could not be done in an automated way. Remarkably, based on the current literature, unlike other eukaryotes, miRNA-like sequences are not conserved among fungi belonging to different genera [70]. Interestingly, three miRNA-like are up-regulated in the intraradical phase, which could lead to the hypothesis of miRNA-like AMF genes required to manipulate fungal or host plant gene expression; however, further analyses are necessary to confirm their possible functional role.

In silico evidences of *M. truncatula* mRNAs potentially targeted by *Rir*-sRNAs

In the absence of a tool specifically designed for the cross-kingdom RNA silencing towards plant transcripts, we have used sPARTA as one of the most powerful tool for target prediction and PARE validation in plants. Our in silico analysis, supported by degradome data, predicted 237 plant genes as putative targets of specific fungal sRNAs. Functional categories associated to these genes shows an enrichment in the GO terms related to

hydrolases and endomembrane systems. Some specific target genes are here discussed in detail for their possible role in the AM symbiosis. The predicted target gene encoding the Specific Tissue (ST) protein 6 (transcript ID: AES73699) belongs to the family of ST proteins, whose function is unknown, but for which transcriptomic data suggest an involvement in biotic and abiotic stress [71, 72]. This protein family has been described in Fabaceae and Asteraceae but seems to be absent in others plant groups such as Brassicaceae [72]. Transcriptomic data suggest their involvement in biotic and abiotic stress [71, 72]. Interestingly, the expression of *M. truncatula ST6* (*MtST6*) gene was found to be modulated during the different steps of the AM association [73, 74] and is induced by fungal diffusible signals, during hyphopodium formation [75] and in arbusculated cells [74]. In this regard, we speculate that one identified *Rir*-sRNA (3121–59) might have a regulatory role on *MtST6* gene expression during the intraradical phase.

Another predicted target is the Responsive To Dehydration 22 (RD22; transcript ID: AES74153) gene, an ABA-dependent signaling gene involved in abiotic stress tolerance [76–78] and in pathogen susceptibility ([79] and references therein). ABA positively regulates AM symbiosis; however, contradictory results have been obtained on ABA content in mycorrhizal roots [80]. In tomato, a gene involved in ABA catabolism (CYP707A3) was specifically expressed in arbuscule-containing cells, while the gene *SINCED*, involved in the ABA biosynthesis, was detected only in cortical cells from non mycorrhizal plants, suggesting that a balance between biosynthesis and catabolism of ABA is determinant for the differentiation of arbuscules [81]. In this context, the targeting of the RD22 gene by three *Rir*-sRNAs (773–218, 10,035–21, 10,035–21) could be related to the down-regulation of this gene in arbusculated cells [74].

Interestingly, among the putative target genes of *Rir*-sRNAs, identified in this study, one encodes a DREPP plasma membrane protein (*MtDREPP*) (transcript ID: KEH37321) which was found to be down-regulated in mycorrhizal roots compared to non-mycorrhizal roots [82]. Host roots undergo plasma membrane (PM) remodeling events during the AMF colonization process from initial contact to intracellular accommodation of fungal structures [83]. In particular, arbuscule accommodation requires both PM expansion and periarbuscular membrane (PAM) generation. These events, that lead to dynamic change of PM protein composition [82] and polarized secretion mediated by exocytotic fusion of membrane vesicles [2], might involve *MtDREPP* modulation.

It is worth mentioning that, among the *Mt*-mRNAs putative targets, we identified the AM-induced *MtVapyrin* (transcript ID: KEH25576) gene (Ankyrin repeat RF-like protein), which is required for arbuscule development [84, 85] and PAM formation [86]. It is tempting

to speculate that its predicted targeting *Rir*-sRNA (2559–70) could contribute to modulate *MtVapyrin* expression in different cell populations and/or during arbuscule formation.

Non-specific phospholipase C4 (*NPC4*) (transcript ID: KEH18078) is another predicted *Rir*-sRNA target gene whose gene product is localized to the PM. It shows high homology with Arabidopsis *NPC4*, a phospholipid-cleaving enzyme responsible for lipid remodeling during phosphate-limiting conditions. In *Arabidopsis*, it has been demonstrated that this gene family could be involved in plant defense response against different pathogens playing a role not only in elicitor recognition processes, but also in downstream disease resistance signaling [87].

Another interesting putative target gene involved in host defense response is a nuclear-binding leucine-rich repeat (NBS-LRR) type disease resistance encoding gene (transcript ID: AES68798) which shows high similarity with rice *OsRGA3*, a *Resistance* (*R*) gene associated with rice blast resistance. Since the response of plants to AMF involves a transient and spatial activation of defense mechanisms [88] the *Rir*-sRNA (7710–27) could be responsible for repressing this gene to allow AMF colonization.

Although supported by computational analyses, further work is needed to experimentally confirm these putative *Rir*-sRNA-*Mtr*-mRNA interactions: 5' RACE assays would be useful to further validate cleavage sites and co-expression of sRNA and its putative mRNA target in transient transformation assay would be helpful to verify the existence of a selective RNAi in vivo [89].

Conclusions

The description of RNAi-related genes, showing an expansion of *AGO*-like and *RdRp* genes, and the characterization of the sRNA population indicate that *R. irregularis* is equipped with a functional sRNA-generating pathway. Our in silico analysis predicted 237 plant genes as putative targets of specific fungal sRNAs suggesting that a cross-kingdom post-transcriptional gene silencing may occur during AMF colonization.

Since HIGS and VIGS tools have been shown to function on AMF, it is likely that interspecies RNA movement also occurs from host plant towards AMF: the dataset generated in this work can be exploited to further investigate plant to fungus RNA exchanges in the AM symbiosis.

Methods

Biological material and growth conditions

All the fungal material (*R. irregularis* DAOM 197198) was obtained from mycorrhizal association with *Medicago truncatula* (Jemalong A17) plants. Nine day old *M. truncatula* seedlings, germinated in sterile conditions, were inoculated, using the Millipore sandwich

method [90], with extraradical fungal structures (ERM) obtained from 2 in vitro monoxenic cultures of *Agrobacterium rhizogenes*-transformed chicory roots in two-compartment Petri plates [91]. In parallel, control non-mycorrhizal plants were treated in the same way, but avoiding the addition of the fungal inoculum. All the plants were fertilized with Long Ashton nutrient solution containing 32 μ M KH₂PO₄ and grown in a climate-controlled room at 22 °C with a photo-period of 14-h light and 10-h dark. After 60 days from the inoculum, plant and fungal materials were harvested. The ERM was manually collected with tweezers under a stereo microscope. Mycorrhizal roots, from which the ERM was removed, were then collected and considered as the intra-radical mycelium (IRM, for qRT-PCR experiment) or mycorrhizal roots (RM, for RNA-seq analysis). Non mycorrhizal roots (RC) were observed under stereomicroscope, to confirm the absence of fungal structures, prior to collection. The harvested material was immediately frozen in liquid nitrogen, lyophilized and stored at -80 °C.

Identification of RdRp, DCL and AGO homologs

We screened, by keywords searches, the recent release of *R. irregularis* genome [40] for the presence of the putative homologs of fungal DCL, RdRp and AGO genes on JGI MycCosm portal [41]. For the identification of AGO sequences, we searched for “piwi” and we retrieved (and considered as AGO-like) all the genes with a Piwi domain (Pfam family: PF02171). Regarding RdRp we retrieved all the genes annotated with a KOG (EuKaryotic Orthologous Groups) ID: KOG0988 (“RNA-directed RNA polymerase QDE-1 required for post-transcriptional gene silencing and RNA interference”). For DCL we searched for “dicer” and we kept only the sequences with two RNase III domains [42].

To identify the RNAi-related homologs genes in *R. clarus*, we first aligned separately AGO-like, RdRp and DCL *R. irregularis* proteins with MAFFT v7.310 (option: --auto) [92] and the alignments were used to build profiles HMM with hmmbuild with default parameters (HMMER 3.1b2 [93]). Then, using hmmsearch (HMMER 3.1b2) with default parameters, we searched for homologs in *R. clarus* proteome [47]. At that point, the resulting sequences were searched for protein domains with hmmscan (options: --cut_ga --domtblout; HMMER 3.1b2) against Pfam-A version 32.0 [94] HMM profiles and then we kept the sequences with a “Piwi” domain for AGO-like, a “RdRP” domain for RdRp and two “Ribonuclease_3” domains for DCL.

Phylogenetic analyses

The whole amino acid sequences of DCL, AGO and RdRp genes were aligned with MAFFT v7.310 (option: --auto) [92]; the alignments were used for phylogenetic inference by the Maximum Likelihood method

implemented in the IQ-TREE software (options: -m TEST -bb 1000 -alrt 1000) [95]. The software performed model selection [96], tree reconstruction and branch support analysis by ultra-fast bootstrap method [97] (1000 replicates). Trees were visualized with Evolview v2 [98]. Protein domains annotations (for tree visualization) were retrieved using hmmscan (options: --cut_ga --domtblout; HMMER 3.1b2) against Pfam-A version 32.0 [94].

In silico gene expression analysis

We retrieved the cDNA of each predicted DCL, AGO-like and RdRp proteins to perform an in silico gene expression analyses, exploiting RNA-seq datasets [40] available at the NCBI Sequence Read Archive (SRR3285893–SRR3285895: 2 day germinating spores; SRR3285917–SRR3285919: symbiotic tissues). Paired reads were trimmed for adapters, filtered for qualities and aligned on cDNA with Bowtie2 (default parameters) [99]. For each sequence we calculated FPKM and we considered expressed all those with a value > 1 (arbitrary selected cut-off) in at least one of the two conditions (germinating spores and symbiotic tissues).

RNA extraction and qRT-PCR assays

Total RNA was extracted with the RNeasy Plant Mini Kit (Qiagen) and then treated with TURBO™ DNase (Ambion). The RNA samples were routinely checked for DNA contamination by PCR analysis, using primers for *MtTef* (RM samples) and for *RiTef* (ERM samples). For cDNA synthesis about 500 ng of total RNA were denatured at 65 °C for 5 min and then reverse-transcribed at 25 °C for 10 min, 42 °C for 50 min and 70 °C for 15 min in a final volume of 20 μ l containing 10 μ M random hexamers, 0.5 mM dNTPs, 4 μ l 5X buffer, 2 μ l 0.1 M DTT and 1 μ l Super-ScriptII (Invitrogen). qRT-PCR experiments were carried out in a final volume of 15 μ l containing 7.5 μ l of iTaq™ Universal SYBR. Green Super-mix (Bio-Rad), 5.5 μ l of 0.8 M primer mix and 2 μ l of 1:10 diluted cDNA. Amplification were run in a Rotor-Gene Q apparatus (Qiagen) using the following program: 5 min pre-incubation at 95 °C and 40 cycles of 30 s at 95 °C, 30 s at 60–64 °C. Each amplification was followed by melting curve analysis (60–94 °C) with a heating rate of 0.5 °C every 15 s. All reactions were performed on at least four biological replicates each with two technical replicates. Relative expression and statistical analyses were performed by REST2009 [100], using as reference genes *RiTef* and *Ri-BetaTubulin1* (for *R. irregularis* RNAi-related gene expression). The presence of a functional AM symbiosis was evaluated (for small RNA-seq experiment) comparing the expression of the *MtPT4* gene relative to the *MtTEF* housekeeping gene in RM (mycorrhizal roots) and RC (control non-mycorrhizal

roots) samples. All primers were previously tested in conventional PCR assays on cDNA, followed by agarose gel electrophoresis, to confirm the specificity and amplification of a single fragment. The list of primers is given in Additional file 9.

RNA extraction for sRNA-seq

For sRNA sequencing, total RNA was extracted with Direct-zol™ RNA MiniPrep (Zymo Research) kit. The concentration and quality of the nucleic acids were assessed with a Nanodrop1000 (Thermo Scientific). Samples were sent to Macrogen (South Korea) for RNA integrity check, library preparations and sequencing. A total of 9 libraries were sequenced: 3 for ERM samples, 3 for RM samples and 3 for RC samples. Each sample was a pool of equal RNA amounts from 3 different biological samples.

Bioinformatics pipeline

Raw sRNA-seq reads, after being checked for quality with FastQC (Babraham Bioinformatics) [101], were cleaned for adapters (TGGAATTCCTCGGGTGCCAA GG), artifacts (default parameters) and low quality reads ($-q\ 28\ -p\ 50$) with Fastx Toolkit (Hannon Lab) [102]. We then removed all the reads mapping on tRNA, rRNA, snRNA and snoRNA on Rfam 12.0 database [103] (Rfam families IDs in Additional file 3) using bowtie aligner [104] allowing up to 1 mismatch. We further filtered reads removing those mapping with 0 mismatch on “ribosomal RNA” sequences of the genus “Rhizophagus” in GenBank and retained only the reads with a length between 18 and 35 nt. Nucleotide length distribution, 5' terminal nucleotide composition and reads redundancy analyses were performed with a set of Perl and R scripts. Reads were mapped on *R. irregularis* DAOM 197198 v2.0 genome on JGI Genome Portal [105] and on *M. truncatula* A17 v4.0 genome on Ensembl [106] with 0 mismatch using bowtie.

Reads mapping on *R. irregularis* genome from ERM and RM libraries were analyzed together in a single run with ShortStack v.3.8.5 [51], for the genome-guided sRNA-generating loci prediction (options: `--mismatches 0 --foldsize 1000 --dicermin 18 --dicermax 35 --pad 200 --mincov 10.0rpm`). The software produced a count table file (with number of reads from each library that defined each locus) that was used for DE analysis between ERM and RM with DESeq2 1.18.1 Bioconductor package [52]. We considered, as differentially expressed, the loci with adjusted *p*-value < 0.05 (Benjamini–Hochberg procedure). ShortStack was also used to produce an annotation file with genomic coordinates of sRNA-generating loci that was used for comparison with *R. irregularis* DAOM 197198 v2.0 gene annotation file with BEDTools [107]. To annotate a *Rir*-sRNA-generating

locus on a specific genomic strand it should originate 80% of reads from the same strand (default parameter in ShortStack). Homology analysis of *Rir*-sRNA-generating loci with fungal repetitive elements from RepBase 23.04 [53] was performed with tblastx [108] (E-value <= 0.005).

For PCA we calculated the nucleotide size proportion of sRNAs for each sRNA-generating locus from 18 to 35 nt (compared to the total of sRNA reads that defined that locus) starting from ShortStack output file and we associated these data with nucleotide length of each locus and total number of reads that defined it. PCA was performed in R with “FactoMineR” (v1.41) package and results visualized with “factoextra” (v1.0.5) package. DBSCAN clustering was performed with fpc package (parameters: `eps = 0.5, minPts = 30`).

MicroRNA-like loci were annotated by ShortStack and their secondary structure were predicted and visualized with StrucVis v.0.3 [109].

To identify *M. truncatula* transcripts potentially targeted by *Rir*-sRNAs in a hypothetical cross-kingdom RNA silencing process, we used sPARTA v.1.20 [56], a software for target prediction and PARE validation previously used for plant datasets, as it is our experimental system for target sequences. We used the 11,396 most expressed *Rir*-sRNAs (merging all the reads from ERM and RM libraries that mapped with 0 mismatch on *R. irregularis* genome) to find targets in *M. truncatula* A17 v.4.0 cDNAs (options: `-tarPred E -tarScore --tag2FASTA --map2DD --validate`). For PARE validation we used published PARE-seq data (SRA accessions: SRR088877, SRR088878) obtained in similar experimental conditions (*M. truncatula*-*R. irregularis* symbiotic association; mycorrhizal and non-mycorrhizal roots; [37]), after being cleaned for adapters and artifacts (default parameters) and filtered for quality ($-q20$) with Fastx Toolkit [102]. The output of sPARTA was filtered to keep only the *Rir*-sRNA-mRNA pairs for which: a) corrected *p*-values < 0.05 (calculated by the software as the confidence score of a sRNA-target interaction corrected for the noise around the cleavage site); b) sRNA length between 21 nt and 24 nt; c) at least 5 PARE reads at cleavage sites from mycorrhizal PARE library; d) no PARE reads at cleavage site from non-mycorrhizal PARE library.

To identify *M. truncatula* transcripts targeted by endogenous sRNAs, we used sPARTA as described above, using as queries the *M. truncatula* miRNAs (*Mtr*-miRNA) from miRBase (Release 22) [58]. The output was filtered to keep only the *Mtr*-miRNA-mRNA pairs for which the following two conditions were both met: a) corrected *p*-values < 0.05; b) at least 5 PARE reads at cleavage sites from mycorrhizal PARE library.

GO enrichment analysis was performed on transcripts identified as potential targets of *Rir*-sRNAs with AgriGO

[110] (p-value < 0.01; statistical test: Fisher's test with Yekutieli multi-test adjustment method).

The analysis of secondary siRNAs production from *Mtr*-mRNAs potentially targeted by *Rir*-sRNAs was performed following the procedure applied by Sahid et al. (2018) [28]. Reads from RM and RC libraries were mapped on *M. truncatula* A17 v.4.0 genome using ShortStack v.3.8.5 (-mismatches 0, -nohp), defining the full length of each mRNA as a locus (option -locifile). The output count table file was used for DE analysis (RM vs RC) with DESeq2 v.1.18.1 as described above. The resulting *Mtr*-mRNAs with an increased number of mapped reads in RM compared to RC were also checked for their presence in the list of potential targets for *Rir*-sRNAs and those in common were analyzed for pahiRNA production with PhaseTank v1.0 [60] (default parameters).

Additional files

Additional file 1: Figure S1. Phylogenetic relationship of AGO proteins in different organisms. **Figure S2.** Phylogenetic relationship of RdRp proteins in different organisms. **Figure S3.** Phylogenetic relationship of AGO, RdRp and DCL proteins in different fungi. **Figure S4.** Expression of *MtPT4* relative to *MTEF* assessed by qRT-PCR in RM samples (mycorrhizal roots) compared to RC ones (nonmycorrhizal roots). Data for each condition are presented as mean \pm standard error. **Figure S5.** Length distribution (expressed in nucleotide) of sRNAs reads (redundant and non-redundant) from RC (non mycorrhizal roots) and RM (mycorrhizal roots) libraries mapping on *Medicago truncatula* genome. **Figure S6.** Relative nucleotide frequency of 5' end of sRNAs reads (redundant and non-redundant) from RM (mycorrhizal roots) and ERM (extra radical mycelium) libraries mapping on *Rhizophagus irregularis* genome. **Figure S7.** Volcano plots (fold changes vs adjusted p-values) of *Rir*-sRNA-generating loci. **Figure S8.** Length distribution (in nucleotide) of sRNA reads that defined the *Rir*-sRNAs-generating loci homologous to repetitive elements in RepBase. Black lines are the length distribution of the individual loci and red line is the average length distribution of the plotted loci. (PDF 3262 kb)

Additional file 2: In silico expression analysis of AGO-like, RdRp and DCL identified in *R. irregularis*. For each sequence we reported: the protein and transcript IDs (JGI), the number of paired mapped reads and FPKM from RNA-seq data-sets (SRR3285893–SRR3285895: 2 day germinating spores; SRR3285917–SRR3285919: symbiotic tissues; Chen et al. 2018) taking into account, the nucleotide length of transcript and the amino-acid length of protein. (XLSX 9 kb)

Additional file 3: Number of sRNA reads after each filtering step and RFAM families IDs (version 12.0) used for rRNA, tRNA, snRNA and snoRNA removal. (XLSX 16 kb)

Additional file 4: Characterization of sRNA-generating loci. For each of the predicted 2131 sRNA-generating loci we reported: the name, the genome coordinates, the miRNA prediction result, the possible overlap with protein-coding genes (reported with JGI protein ID), the differential expression analysis results of differential expression analysis, the length, the number of mapped sRNA reads, the eventual ID of homologous loci from RepBase (when available), the results of DBSCAN clustering, the proportion of mapped sRNA reads of each length (from 18 nt to 35 nt). (XLSX 442 kb)

Additional file 5: Results of *Rir*-sRNA – *Mtr*-mRNA target prediction performed with sPARTA. We added at the filtered software output file the information about the number of sRNA reads, the length of sRNAs (nt), the description of targets and the names of the sRNA-generating loci that produce the sRNAs. (XLSX 47 kb)

Additional file 6: Results of *Mtr*-miRNA – *Mtr*-mRNA target prediction performed with sPARTA. We added the description of targets to the filtered software output file. (XLSX 35 kb)

Additional file 7: Prediction results for *Mtr*-mRNAs that are targeted by both *Rir*-sRNAs and *Mtr*-miRNAs according to sPARTA analysis. (XLSX 9 kb)

Additional file 8: Results of DE analysis of sRNA reads originated from *Mtr*-mRNAs for identification of mycorrhizal induced secondary siRNAs. We reported only the transcripts with a significant increase in number of mapped sRNAs reads in RM (mycorrhizal roots) compared to RC (non-mycorrhizal roots) condition (log2FoldChange > 0, padj < 0.05; DESeq2). (XLSX 55 kb)

Additional file 9: List of oligonucleotide used in qRT-PCR assays. (XLSX 46 kb)

Abbreviations

AMF: Arbuscular Mycorrhizal Fungi; ERM: Extra radical mycelium; GO: Gene Ontology; RC: Non-mycorrhizal roots; RM: Mycorrhizal roots; sRNA: Small RNA

Acknowledgments

Not Applicable.

Funding

Research work was supported by the Italian Ministry for University and Research (MIUR - UNITO Ricerca Locale 2016) to L.L. The funding body was not involved in the design of the study or in any aspect of the collection, analysis and interpretation of the data or paper writing.

Availability of data and materials

Raw small RNA-seq data are available in European Nucleotide Archive (ENA). Study = ERP111459; Accession = ERR2841782, ERR2841783, ERR2841784 (Extra radical mycelium); ERR2841785, ERR2841786, ERR2841787 (Mycorrhizal root); ERR2841788, ERR2841789, ERR2841790 (Non mycorrhizal roots).

Authors' contributions

LL, MT and GPA designed and coordinated the work. AS and VF prepared the biological material, extracted the nucleic acids and performed the qRT-PCR assays. AS performed the characterization of RNAi-related genes and, together with LM, carried out to the computational analyses of sRNAs. AS together with LL and MT drafted the manuscript. All authors contributed to the assessment of results, manuscript revisions and final approval for publication.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors declare consent to publish.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Life Sciences and Systems Biology, University of Torino, Viale P.A. Mattioli 25, 10125 Torino, Italy. ²Institute for Sustainable Plant Protection – CNR Torino, Strada delle Cacce 73, 10131 Torino, Italy.

Received: 29 October 2018 Accepted: 22 February 2019

Published online: 04 March 2019

References

- Spatafora JW, Chang Y, Benny GL, Lazarus K, Smith ME, Berbee ML, et al. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia*. 2016;108:1028–46. <https://doi.org/10.3852/16-042>.
- MacLean AM, Bravo A, Harrison MJ. Plant signaling and metabolic pathways enabling arbuscular mycorrhizal symbiosis. *Plant Cell*. 2017; tpc.00555.2017. <https://doi.org/10.1105/tpc.17.00555>.
- Jacott C, Murray J, Ridout C. Trade-offs in arbuscular mycorrhizal symbiosis: disease resistance, growth responses and perspectives for crop breeding. *Agronomy*. 2017;7:75. <https://doi.org/10.3390/agronomy7040075>.

4. Lanfranco L, Fiorilli V, Gutjahr C. Partner communication and role of nutrients in the arbuscular mycorrhizal symbiosis. *New Phytol.* 2018;220(4): 1031–46. <https://doi.org/10.1111/nph.15230>.
5. Guimil S, Chang H-S, Zhu T, Sesma A, Osbourn A, Roux C, et al. Comparative transcriptomics of rice reveals an ancient pattern of response to microbial colonization. *Proc Natl Acad Sci.* 2005;102:8066–70. <https://doi.org/10.1073/pnas.0502999102>.
6. Liu J, Maldonado-Mendoza I, Lopez-Meyer M, Cheung F, Town CD, Harrison MJ. Arbuscular mycorrhizal symbiosis is accompanied by local and systemic alterations in gene expression and an increase in disease resistance in the shoots. *Plant J.* 2007;50:529–44. <https://doi.org/10.1111/j.1365-313X.2007.03069.x>.
7. Guether M, Balestrini R, Hannah M, He J, Udvardi MK, Bonfante P. Genome-wide reprogramming of regulatory networks, transport, cell wall and membrane biogenesis during arbuscular mycorrhizal symbiosis in *Lotus japonicus*. *New Phytol.* 2009;182:200–12. <https://doi.org/10.1111/j.1469-8137.2008.02725.x>.
8. Zouari I, Salvioi A, Chialva M, Novero M, Miozzi L, Tenore GC, et al. From root to fruit: RNA-Seq analysis shows that arbuscular mycorrhizal symbiosis may affect tomato fruit metabolism. *BMC Genomics.* 2014;15:221. <https://doi.org/10.1186/1471-2164-15-221>.
9. Fiorilli V, Vannini C, Ortolani F, Garcia-Seco D, Chiappello M, Novero M, et al. Omics approaches revealed how arbuscular mycorrhizal symbiosis enhances yield and resistance to leaf pathogen in wheat. *Sci Rep.* 2018;8:9625. <https://doi.org/10.1038/s41598-018-27622-8>.
10. Tisserant E, Kohler A, Dozolme-Seddas P, Balestrini R, Benabdellah K, Colard A, et al. The transcriptome of the arbuscular mycorrhizal fungus *Glomus intraradices* (DAOM 197198) reveals functional trade-offs in an obligate symbiont. *New Phytol.* 2012;193:755–69.
11. Wilson RC, Doudna JA. Molecular mechanisms of RNA interference. *Annu Rev Biophys.* 2013;42:217–39. <https://doi.org/10.1146/annurev-biophys-083012-130404>.
12. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nat Rev Genetics.* 2009;10:94–108. <https://doi.org/10.1038/nrg2504>.
13. Torres-Martinez S, Ruiz-Vázquez RM. The RNAi universe in fungi: a varied landscape of small RNAs and biological functions. *Annu Rev Microbiol.* 2017;71:371–91. <https://doi.org/10.1146/annurev-micro-090816-093352>.
14. Romano N, Macino G. Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Mol Microbiol.* 1992;6:3343–53. <https://doi.org/10.1111/j.1365-2958.1992.tb02202.x>.
15. Nunes CC, Sailsbery JK, Dean RA. Characterization and application of small RNAs and RNA silencing mechanisms in fungi. *Fungal Biol Rev.* 2011;25:172–80. <https://doi.org/10.1016/j.fbr.2011.10.001>.
16. Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, et al. RNAi in budding yeast. *Science.* 2009;326:544–50. <https://doi.org/10.1126/science.1176945>.
17. Knip M, Constantin ME, Thordal-Christensen H. Trans-kingdom cross-talk: small RNAs on the move. *PLoS Genet.* 2014;10(9):e1004602. <https://doi.org/10.1371/journal.pgen.1004602>.
18. Wang M, Weiberg A, Lin F-M, Thomma BPHJ, Huang H-D, Jin H. Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. *Nat Plants.* 2016;2:16151. <https://doi.org/10.1038/nplants.2016.151>.
19. Weiberg A, Wang M, Bellinger M, Jin H. Small RNAs: a new paradigm in plant-microbe interactions. *Annu Rev Phytopathol.* 2014;52:495–516. <https://doi.org/10.1146/annurev-phyto-102313-045933>.
20. Weiberg A, Jin H. Small RNAs—the secret agents in the plant-pathogen interactions. *Curr Opin Plant Biol.* 2015;26:87–94. <https://doi.org/10.1016/j.pbi.2015.05.033>.
21. Weiberg A, Bellinger M, Jin H. Conversations between kingdoms: small RNAs. *Curr Opin Biotechnol.* 2015;32:207–15. <https://doi.org/10.1016/j.copbio.2014.12.025>.
22. Zhou G, Zhou Y, Chen X. New insight into inter-kingdom communication: horizontal transfer of mobile small RNAs. *Front Microbiol.* 2017;8:1–9.
23. Wang M, Thomas N, Jin H. Cross-kingdom RNA trafficking and environmental RNAi for powerful innovative pre- and post-harvest plant protection. *Curr Opin Plant Biol.* 2017;38:133–41. <https://doi.org/10.1016/j.pbi.2017.05.003>.
24. Cai Q, He B, Kogel K-H, Jin H. Cross-kingdom RNA trafficking and environmental RNAi — nature's blueprint for modern crop protection strategies. *Curr Opin Microbiol.* 2018;46:58–64. <https://doi.org/10.1016/j.mib.2018.02.003>.
25. Chaloner T, van Kan JAL, Grant-Downton RT. RNA “information warfare” in pathogenic and mutualistic interactions. *Trends Plant Sci.* 2016;21:738–48.
26. Weiberg A, Wang M, Lin FM, Zhao H, Zhang Z, Kaloshian I, et al. Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science.* 2013;342:118–23. <https://doi.org/10.1126/science.1239705>.
27. Cai Q, Qiao L, Wang M, He B, Lin F-M, Palmquist J. Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. *Science.* 2018;360(6393):1126–9. <https://doi.org/10.1126/science.aar4142>.
28. Shahid S, Kim G, Johnson NR, Wafala E, Wang F, Coruh C, et al. MicroRNAs from the parasitic plant *Cuscuta campestris* target host messenger RNAs. *Nature.* 2018;553:82–5. <https://doi.org/10.1038/nature25027>.
29. Wang M, Jin H. Spray-induced gene silencing: a powerful innovative strategy for crop protection. *Trends Microbiol.* 2017;25:4–6. <https://doi.org/10.1016/j.tim.2016.11.011>.
30. Lee SJ, Kong M, Harrison P, Hijiri M. Conserved proteins of the RNA interference system in the arbuscular mycorrhizal fungus *Rhizoglyphus irregularis* provide new insight into the evolutionary history of glomeromycota. *Genome Biol Evol.* 2018;10:328–43. <https://doi.org/10.1093/gbe/evy002>.
31. Kikuchi Y, Hijikata N, Ohtomo R, Handa Y, Kawaguchi M, Saito K, et al. Aquaporin-mediated long-distance polyphosphate translocation directed towards the host in arbuscular mycorrhizal symbiosis: application of virus-induced gene silencing. *New Phytol.* 2016;211:1:202–8.
32. Helber N, Wipfel K, Sauer N, Schaarschmidt S, Hause B, Requena N. A versatile monosaccharide transporter that operates in the arbuscular mycorrhizal fungus *Glomus* sp is crucial for the symbiotic relationship with plants. *Plant Cell.* 2011;23:3812–23. <https://doi.org/10.1105/tpc.111.089813>.
33. Tsuzuki S, Handa Y, Takeda N, Kawaguchi M. Strigolactone-induced putative secreted protein 1 is required for the establishment of symbiosis by the arbuscular mycorrhizal fungus *Rhizophagus irregularis*. *Mol Plant-Microbe Interact.* 2016;29:1–59.
34. Voß S, Betz R, Heidt S, Corradi N, Requena N. RICRN1, a crinkler effector from the arbuscular mycorrhizal fungus *Rhizophagus irregularis*, functions in arbuscule development. *Front Microbiol.* 2018;9:1–18.
35. Formey D, Sallet E, Lelandais-Brière C, Ben C, Bustos-Sanmamed P, Niebel A, et al. The small RNA diversity from *Medicago truncatula* roots under biotic interactions evidences the environmental plasticity of the miRNAome. *Genome Biol.* 2014;15:457. <https://doi.org/10.1186/s13059-014-0457-4>.
36. Lelandais-Brière C, Moreau J, Hartmann C, Crespi M. Noncoding RNAs, emerging regulators in root endosymbioses. *Mol Plant-Microbe Interact.* 2016;29:170–80. <https://doi.org/10.1094/MPMI-10-15-0240-F1>.
37. Devers EA, Branscheid A, May P, Krajinski F. Stars and Symbiosis: MicroRNA- and microRNA*-mediated transcript cleavage involved in arbuscular mycorrhizal symbiosis. *Plant Physiol.* 2011;156:1990–2010. <https://doi.org/10.1104/pp.111.172627>.
38. Wu P, Wu Y, Liu C-C, Liu L-W, Ma F-F, Wu X-Y, et al. Identification of arbuscular mycorrhiza (AM)-responsive microRNAs in tomato. *Front Plant Sci.* 2016;7:429. <https://doi.org/10.3389/fpls.2016.00429>.
39. Couzigou JM, Laressergues D, André O, Gutjahr C, Guillotin B, Bécard G, et al. Positive gene regulation by a natural protective miRNA enables arbuscular mycorrhizal symbiosis. *Cell Host Microbe.* 2017;21:106–12. <https://doi.org/10.1016/j.chom.2016.12.001>.
40. Chen ECH, Morin E, Beaudet D, Noel J, Yildirim G, Ndikumana S, et al. High intraspecific genome diversity in the model arbuscular mycorrhizal symbiont *Rhizophagus irregularis*. *New Phytol.* 2018;220(4):1161–71. <https://doi.org/10.1111/nph.14989>.
41. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otilar R, et al. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 2014;42: D699–704. <https://doi.org/10.1093/nar/gkt1183>.
42. MacRae IJ. Structural basis for double-stranded RNA processing by dicer. *Science.* 2006;311:195–8. <https://doi.org/10.1126/science.1121638>.
43. Iyer LM, Koonin EV, Aravind L. Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. *BMC Struct Biol.* 2003;3:1–23.
44. Torres-Martinez S, Ruiz-Vázquez RM. RNAi pathways in Mucor: a tale of proteins, small RNAs and functional diversity. *Fungal Genet Biol.* 2016;90:44–52. <https://doi.org/10.1016/j.fgb.2015.11.006>.
45. Poulsen C, Vaucheret H, Brodersen P. Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. *Plant Cell.* 2013;25:22–37. <https://doi.org/10.1105/tpc.112.105643>.

46. Hutvagner G, Simard MJ. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol.* 2008;9:22–32. <https://doi.org/10.1038/nrm2321>.
47. Kobayashi Y, Maeda T, Yamaguchi K, Kameoka H, Tanaka S, Ezawa T, et al. The genome of *Rhizophagus clarus* HR1 reveals a common genetic basis for autotrophy among arbuscular mycorrhizal fungi. *BMC Genomics.* 2018;19:465. <https://doi.org/10.1186/s12864-018-4853-0>.
48. Mueth NA, Ramachandran SR, Hulbert SH. Small RNAs from the wheat stripe rust fungus (*Puccinia striiformis* f.sp. *tritica*). *BMC Genomics.* 2015;16:718. <https://doi.org/10.1186/s12864-015-1895-4>.
49. Lelandais-Brière C, Naya L, Sallet E, Calenge F, Frugier F, Hartmann C, et al. Genome-wide *Medicago truncatula* small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell.* 2009;21:2780–96. <https://doi.org/10.1105/tpc.109.068130>.
50. Borges F, Martienssen RA. The expanding world of small RNAs in plants. *Nat Publ Gr.* 2015;16:1–15. <https://doi.org/10.1038/nrm4085>.
51. Johnson NR, Yeoh JM, Coruh C, Axtell MJ. Improved Placement of Multi-mapping Small RNAs. *G3 (Bethesda).* 2016;6:2103–11. <https://doi.org/10.1534/g3.116.030452>.
52. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
53. Repbase - GIRI. <https://www.girinst.org/repbase/>. Accessed 26 Oct 2018.
54. Fahlgren N, Bollmann SR, Kasschau KD, Cuperus JT, Press CM, Sullivan CM, et al. *Phytophthora* have distinct endogenous small RNA populations that include short interfering and microRNAs. *PLoS One.* 2013;8:e77181. <https://doi.org/10.1371/journal.pone.0077181>.
55. Daszykowski M, Walczak B. Density-Based clustering methods. *Comprehensive Chemometrics.* 2010:635–54. <https://doi.org/10.1016/B978-0-44452701-1.00067-3>.
56. Kakrana A, Hammond R, Patel P, Nakano M, Meyers BC. SPARTA: a parallelized pipeline for integrated analysis of plant miRNA and cleaved mRNA data sets, including new miRNA target-identification software. *Nucleic Acids Res.* 2014;42:e139. <https://doi.org/10.1093/nar/gku693>.
57. Axtell MJ. Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol.* 2013;64:137–59. <https://doi.org/10.1146/annurev-arplant-050312-120043>.
58. Kozomara A, Griffiths-Jones S. MiRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 2014;42:D68–73. <https://doi.org/10.1093/nar/gkt1181>.
59. Fei Q, Xia R, Meyers BC. Phased, secondary, small interfering RNAs in post transcriptional regulatory networks. *Plant Cell.* 2013;25:2400–15. <https://doi.org/10.1105/tpc.113.114652>.
60. Guo Q, Qu X, Jin W. PhaseTank: genome-wide computational identification of phasiRNAs and their regulatory cascades. *Bioinformatics.* 2015;31:284–6. <https://doi.org/10.1093/bioinformatics/btu628>.
61. Kamel L, Keller-Pearson M, Roux C, Ané J-M. Biology and evolution of arbuscular mycorrhizal symbiosis in the light of genomics. *New Phytol.* 2016;57–64.
62. Tisserant E, Malbreil M, Kuo A, Kohler A, Symeonidi A, Balestrini R, et al. Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proc Natl Acad Sci U S A.* 2013;110:20117–22. <https://doi.org/10.1073/pnas.1313452110>.
63. Lin K, Limpens E, Zhang Z, Ivanov S, Saunders DGO, Mu D, et al. Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. *PLoS Genet.* 2014;10(1):e1004078. <https://doi.org/10.1371/journal.pgen.1004078>.
64. Chang S-S, Zhang Z, Liu Y. RNA interference pathways in fungi: mechanisms and functions. *Annu Rev Microbiol.* 2012;66:305–23. <https://doi.org/10.1146/annurev-micro-092611-150138>.
65. Swarts DC, Jore MM, Westra ER, Zhu Y, Janssen JH, Snijders AP, et al. DNA-guided DNA interference by a prokaryotic Argonaute. *Nature.* 2014;507:258–61. <https://doi.org/10.1038/nature12971>.
66. Raman V, Simon SA, Demirci F, Nakano M, Meyers BC, Donofrio NM. Small RNA functions are required for growth and development of *Magnaporthe oryzae*. *Mol Plant-Microbe Interact.* 2017;30:517–30. <https://doi.org/10.1094/MPMI-11-16-0236-R>.
67. Trieu TA, Calo S, Nicolás FE, Vila A, Moxon S, Dalmay T, et al. A non-canonical RNA silencing pathway promotes mRNA degradation in basal fungi. *PLoS Genet.* 2015;11:1–32.
68. Cervantes M, Vila A, Nicolás FE, Moxon S, de Haro JP, Dalmay T, et al. A single argonaute gene participates in exogenous and endogenous RNAi and controls cellular functions in the basal fungus *Mucor circinelloides*. *PLoS One.* 2013;8(7):e69283. <https://doi.org/10.1371/journal.pone.0069283>.
69. Nicolas FE, Moxon S, de Haro JP, Calo S, Grigoriev IV, Torres-Martínez S, et al. Endogenous short RNAs generated by dicer 2 and RNA-dependent RNA polymerase 1 regulate mRNAs in the basal fungus *Mucor circinelloides*. *Nucleic Acids Res.* 2010;38:5535–41.
70. Villalobos-Escobedo JM, Herrera-Estrella A, Carreras-Villaseña R. The interaction of fungi with the environment orchestrated by RNAi. *Mycologia.* 2016;108:556–71. <https://doi.org/10.3852/15-246>.
71. Albornos L, Martín I, Hernández-Nistal J, Labrador E, Dopico B. Three members of *Medicago truncatula* ST family (MtST4, MtST5 and MtST6) are specifically induced by hormones involved in biotic interactions. *Plant Physiol Biochem.* 2018;127:496–505.
72. Albornos L, Martín I, Iglesias R, Jiménez T, Labrador E, Dopico B. ST proteins, a new family of plant tandem repeat proteins with a DUF2775 domain mainly found in Fabaceae and Asteraceae. *BMC Plant Biol.* 2012;12:207. <https://doi.org/10.1186/1471-2229-12-207>.
73. Hohnjec N. 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 Physiol.* 2005;137:1283–301. <https://doi.org/10.1104/pp.104.056572>.
74. Gaudé N, Bortfeld S, Duensing N, Lohse M, Krajinski F. Arbuscule-containing and non-colonized cortical cells of mycorrhizal roots undergo extensive and specific reprogramming during arbuscular mycorrhizal development. *Plant J.* 2012;69:510–28.
75. Kuhn H, Küster H, Requena N. Membrane steroid-binding protein 1 induced by a diffusible fungal signal is critical for mycorrhization in *Medicago truncatula*. *New Phytol.* 2010;185:716–33.
76. Iwasaki T, Yamaguchi-Shinozaki K, Shinozaki K. Identification of a cis-regulatory region of a gene in *Arabidopsis thaliana* whose induction by dehydration is mediated by abscisic acid and requires protein synthesis. *Mol Gen Genet.* 1995;247:391–8. <https://doi.org/10.1007/BF00293139>.
77. Abe H. Role of Arabidopsis MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. *Plant Cell.* 1997;9:1859–68. <https://doi.org/10.1105/tpc.9.10.1859>.
78. Abe H. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell.* 2003;15:63–78. <https://doi.org/10.1105/tpc.006130>.
79. Matus JT, Aquea F, Espinoza C, Vega A, Cavallini E, Dal Santo S, et al. Inspection of the grapevine BURP superfamily highlights an expansion of RD22 genes with distinctive expression features in berry development and ABA-mediated stress responses. *PLoS One.* 2014;9(10):e110372. <https://doi.org/10.1371/journal.pone.0110372>.
80. Herrera-Medina MJ, Steinkellner S, Vierheilig H, Ocampo Bote JA, García Garrido JM. Abscisic acid determines arbuscule development and functionality in the tomato arbuscular mycorrhiza. *New Phytol.* 2007;175:554–64.
81. Fiorilli V, Catoni M, Miozzi L, Novero M, Accotto GP, Lanfranco L. Global and cell-type gene expression profiles in tomato plants colonized by an arbuscular mycorrhizal fungus. *New Phytol.* 2009;184:975–87. <https://doi.org/10.1111/j.1469-8137.2009.03031.x>.
82. Aloui A, Recorbet G, Lemaître-Guillier C, Mounier A, Balliau T, Zivy M, et al. The plasma membrane proteome of *Medicago truncatula* roots as modified by arbuscular mycorrhizal symbiosis. *Mycorrhiza.* 2018;28:1–16.
83. Harrison MJ, Ivanov S. Exocytosis for endosymbiosis: membrane trafficking pathways for development of symbiotic membrane compartments. *Curr Opin Plant Biol.* 2017;38:101–8. <https://doi.org/10.1016/j.cup.2017.04.019>.
84. Feddermann N, Reinhardt D. Conserved residues in the ankyrin domain of VAPYRIN indicate potential protein-protein interaction surfaces. *Plant Signal Behav.* 2011;6:680–4.
85. Pumplin N, Mondo SJ, Topp S, Starker CG, Gantt JS, Harrison MJ. *Medicago truncatula* Vapyrin is a novel protein required for arbuscular mycorrhizal symbiosis. *Plant J.* 2010;61:482–94.
86. Zhang X, Pumplin N, Ivanov S, Harrison MJ. EXO70L is required for development of a sub-domain of the periarbuscular membrane during arbuscular mycorrhizal symbiosis. *Curr Biol.* 2015;25:2189–95. <https://doi.org/10.1016/j.cub.2015.06.075>.
87. Siebers M, Brands M, Wewer V, Duan Y, Hölzl G, Dörmann P. Lipids in plant-microbe interactions. *Biochim Biophys Acta - Mol Cell Biol Lipids.* 1861;2016:1379–95. <https://doi.org/10.1016/j.bbalip.2016.02.021>.
88. García-Garrido JM, Ocampo JA. Regulation of the plant defence response in arbuscular mycorrhizal symbiosis. *J Exp Bot.* 2002;53:1377–86. <https://doi.org/10.1093/jxb/53.373.1377>.
89. Wang M, Weiberg A, Dellota E, Yamane D, Jin H. *Botrytis* small RNA Bc-siR37 suppresses plant defense genes by cross-kingdom RNAi. *RNA Biol.* 2017;14:421–8.

90. Giovannetti M, Sbrana C, Avio L, Citernes AS, Logi C. Differential hyphal morphogenesis in arbuscular mycorrhizal fungi during pre-infection stages. *New Phytol.* 1993;125:587–93.
91. Belmondo S, Fiorilli V, Pérez-Tienda J, Ferrol N, Marmeisse R, Lanfranco L. A dipeptide transporter from the arbuscular mycorrhizal fungus *Rhizophagus irregularis* is upregulated in the intraradical phase. *Front Plant Sci.* 2014;4:36. <https://doi.org/10.3389/fpls.2014.00436>.
92. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013; 30:772–80. <https://doi.org/10.1093/molbev/mst010>.
93. hmmer. <http://hmmer.org/>. Accessed 26 Oct 2018.
94. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. *Nucleic Acids Res.* 2014;42:290–301.
95. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 2015;32:268–74.
96. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 2017;14:587–91.
97. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol.* 2018;35: 518–22. <https://doi.org/10.1093/molbev/msx281>.
98. He Z, Zhang H, Gao S, Lercher MJ, Chen WH, Hu S. Evolvew v2: an online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res.* 2016;44:W236–41. <https://doi.org/10.1093/nar/gkw370>.
99. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods.* 2012;9:357–9.
100. Pfaffl MW, Horgan GW, Dempfle L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.* 2002;30:e36. <https://doi.org/10.1093/nar/30.9.e36>.
101. FastQC. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 26 Oct 2018.
102. Fastx Toolkit. http://hannonlab.cshl.edu/fastx_toolkit/index.html. Accessed 26 Oct 2018.
103. Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, et al. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* 2015;43:D130–7.
104. Langmead B, Trapnell C, Pop M, Salzberg S. 2C- ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10:R25.
105. Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, et al. The genome portal of the department of Energy Joint Genome Institute. *Nucleic Acids Res.* 2012;40:D26–32. <https://doi.org/10.1093/nar/gkr947>.
106. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. *Nucleic Acids Res.* 2018;46:D754–61. <https://doi.org/10.1093/nar/gkx1098>.
107. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26:841–2.
108. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10:421. <https://doi.org/10.1186/1471-2105-10-421>.
109. strucVis. <https://github.com/MikeAxtell/strucVis>. Accessed 26 Oct 2018.
110. Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, et al. AgriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 2017; 45:W122–9. <https://doi.org/10.1093/nar/gkx382>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



CHAPTER 3

3. *Rhizophagus irregularis* sRNA *Rir-2216* is likely involved in cross-kingdom RNAi by silencing host plant gene

3.1. Introduction

Experimental data generated in recent years suggest that cross-kingdom RNAi-like mechanisms occur in AMS. In particular, as described in the previous chapters, Host-Induced (HIGS) and Virus-Induced Gene Silencing (VIGS) have been successfully applied to perform gene knock-down in AMF, through the expression *in planta* of silencing constructs directed against fungal genes (Helber et al., 2011; Kikuchi et al., 2016; Tsuzuki et al., 2016; Xie et al., 2016; Voß et al., 2018). Moreover, two recent reports showed extensive membranes rearrangements and extracellular vesicles (EVs) formation at plant-AMF interface (Ivanov et al., 2019; Roth et al., 2019). Considering that EVs seem to be the principal system through which RNAs move between interacting organisms in cross-kingdom RNAi interactions (Huang et al., 2019), these observations further support the hypothesis envisaging active sRNA exchange between partners in AMS.

In a previous work, we characterized the sRNAome of the AMF *Rhizophagus irregularis* during the symbiosis with the host plant *Medicago truncatula* and we reported *in silico* evidence of cross-kingdom RNAi (Silvestri et al., 2019). The output of bioinformatic analyses are predictions, and therefore some validations through wet lab experiments are needed.

Here, we set up a number of assays to provide a validation of the results of the sRNA analyses and to offer further experimental evidence that fungal sRNAs could guide the silencing of plant genes in the AMS. We first provide evidence, by means of RT-PCR assays, of the existence of a selection of sRNAs identified through sRNA-seq. We then exploited the same sRNA-seq dataset to identify, through a different *in silico* approach, all the fungal sRNAs up-regulated in the intraradical phase that are potentially able to silence plant transcripts. We experimentally validated some of these predictions (fungal sRNA-plant mRNA target pairs) by means of transient co-expression assays in *Nicotiana benthamiana* leaves. Moreover, we also investigated the expression profiles of putative plant target genes at a cell-type level thanks to the use of the laser microdissection (LMD) coupled with qRT-PCR.

In their whole the data we obtained support that some fungal sRNA can interfere with plant gene expression.

3.2. Results and Discussion

Confirmation of sRNA-seq results by RT-PCR

As a first step we wanted to validate the results from a previous sRNA-seq experiment from mycorrhizal roots (Silvestri et al. 2019) and demonstrate the existence of some selected fungal and plant sRNA with an independent experimental approach. A RT-PCR assay for sRNA amplification was therefore set up (Varkonyi-Gasic et al., 2007).

We focused on 2 *R. irregularis* sRNAs (*Rir-196*: 5'-CCCGAGGCTGGAAACAGAGCAGG-3'; and *Rir-434*: 5'-TAATGATGAGACTTGTGGGAG-3') with a relatively high expression level (about 122 and 50 RPM - Reads Per Million Reads - respectively) and 1 *M. truncatula* miRNA (*Mtr-mir399a*: 5'-TGCCAAAGGAGATTTGCCAG-3'). *Mtr-mir399a* belongs to the conserved mir399 family, known to play a role in plant phosphate-starvation response (Chiou et al., 2006). Moreover, *Mtr-mir399a* was found to be expressed at high levels in arbuscules (Branscheid et al., 2010).

Amplification products of the expected size were observed for *Rir-196* (Fig. 1A) and *Rir-434* (Fig. 1B) in RM (mycorrhizal roots) and ERM (extraradical mycelium) samples but not in RC (control non mycorrhizal roots) samples. The plant *Mtr-mir399a* was successfully amplified in RC and RM samples while a PCR product, but of weak intensity, was also observed in the ERM sample (Fig. 1C); this is probably a consequence of a contamination of ERM samples with root material during sample harvesting, as it was already noticed from sRNA-seq mapping results (Silvestri et al., 2019).

These experiments allowed us to confirm the validity of the sRNA-seq approach (Silvestri et al., 2019): we indeed demonstrated that the sRNA reads assigned to *R. irregularis* originate from the fungus and that the one assigned to *M. truncatula* originates from the plant

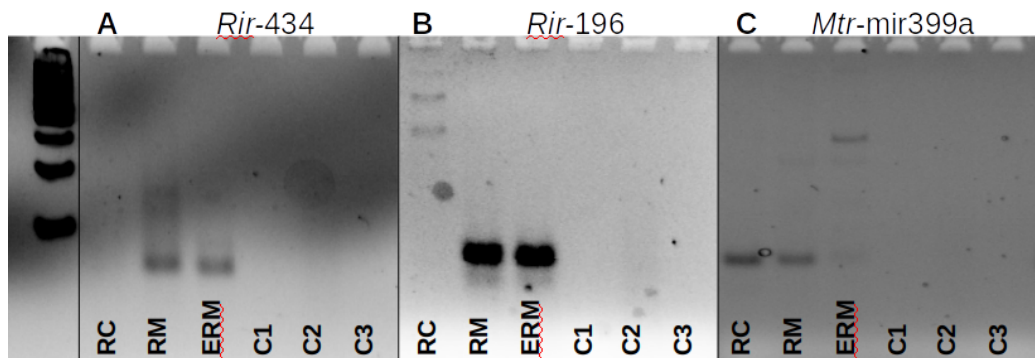


Figure 1. RT-PCR to detect two *Rhizophagus irregularis* sRNAs (*Rir-434*, **A**; and *Rir-196*, **B**) and a *Medicago truncatula* miRNA (*Mtr-mir399a*, **C**) in different samples (RC = control non-mycorrhizal roots; RM = mycorrhizal roots; ERM = extra-radical mycelium; C1 = no RNA RT control; C2 = no enzyme RT control; C3 = H₂O PCR control). The first lane corresponds to 100 bp DNA ladder (first band = 100 bp). All the PCR products are smaller than 100 bp, as expected.

***R. irregularis* sRNA can target several *M. truncatula* transcripts according to *in silico* analysis: the case of *Rir-2216*.**

In order to identify *M. truncatula* transcripts (*Mtr*-mRNAs) potentially targeted by *R. irregularis* sRNAs (*Rir*-sRNAs), we followed a different *in silico* approach from the one previously described, exploiting the same sRNA-seq data (Silvestri et al., 2019). In this case we only used the most abundant fungal sRNAs from each *R. irregularis* sRNA-generating locus up-regulated in the intra-radical (RM) compared to the extra-radical (ERM) conditions to perform a target prediction against plant transcriptome with psRNAtarget (Expectation ≤ 2.5 ; Dai, et al., 2017). This analysis identified, under these stringent parameters, 2355 potential *Mtr*-mRNA targets for 415 *Rir*-sRNAs (Suppl. File 1).

We chose to further study the fungal sRNA *Rir-2216* (5'-TTTTCTTCATCTTCCTCG-3') considering it can potentially target 246 plant transcripts that, for the conservative approach we used, is a very high number. *Rir-2216* is encoded by the sRNA-generating locus 832 and its expression level in sRNA-seq libraries is about 9 RPM. We designed specific oligonucleotides to validate *Rir-2216* by means of RT-PCR. *Rir-2216* was successfully amplified in RM and ERM samples leading a PCR product of the expected size; but a PCR product of weak intensity and smaller in size

compared to that obtained from RM and ERM was observed in RC samples (Fig. 2), probably originated from an aspecific amplification. This PCR product was obtained even increasing the PCR annealing-extension temperature up to 66.1°C (Fig. 2). Sequencing of this DNA band will allow to define its precise sequence and origin. Anyway, as no sRNA reads corresponding to *Rir-2216* were observed in the RC sRNA-seq libraries, we considered *Rir-2216* as a *bonafide* fungal sRNA.

Rir-2216

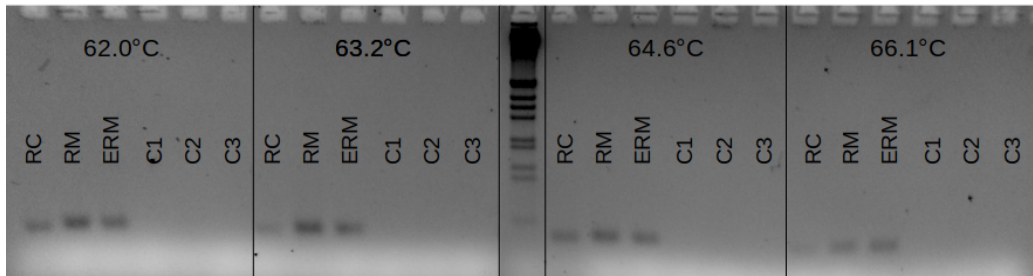


Figure 2. RT-PCR to detect two *Rhizophagus irregularis* sRNAs *Rir-2216* in different samples (RC = control non-mycorrhizal roots; RM = mycorrhizal roots; ERM = extra-radical mycelium; C1 = no RNA RT control; C2 = no enzyme RT control; C3 = H₂O PCR control) and at different annealing-extension temperatures. The central lane corresponds to 100 bp DNA ladder (first band = 100 bp). All the PCR products are smaller than 100 bp, as expected.

***Rir-2216* silences two of its predicted plant targets in transient co-expression assay**

In order to validate the *in silico* target prediction analysis for *Rir-2216*, we performed an *Agrobacterium tumefaciens*-mediated transient co-expression assay in *Nicotiana benthamiana* leaves (Weiberg et al., 2013; Wang et al., 2017) with two of its predicted *M. truncatula* target genes: a transcription factor belonging to the WRKY family (*Mtr-WRKY*; AES66958) and a Lipid transfer protein (*Mtr-LTP*; AES87505). Even if a specific characterization on these 2 genes is missing, it is known that WRKY are involved in pathogen-related stress responses (Hahlbrock et al., 2003) as well as LTP are involved in generation of systemic signaling in response to pathogens (Maldonado et al., 2002). Moreover, some WRKY are also known to be induced upon AMF root colonization (Gaude et al., 2011; Hoge Kamp et al., 2011).

Mtr-WRKY is targeted by *Rir-2216* at its coding sequence (CDS) while *Mtr-LTP* at its 3'UTR, both with a nearly perfect complementarity (Expectation = 1; Fig.

3A). We first cloned *Rir-2216* into a plant artificial miRNA vector (*amiR-Rir-2216*) and then into the pEarlygate 100 (pE100) binary expression vector. At the same time, we cloned the whole *Mtr-WRKY* cDNA sequence into the pEarlygate101 (pE101) binary expression vector characterized by hemagglutinin (HA) and a YFP tag at the C-terminus (*Mtr-WRKY-YFP-HA*). For as regards *Mtr-LTP*, we only cloned its 21-nt long target sequence (5'-TGAGTAGGAAGATGAAGAAAA-3') at the 3'-end of YFP CDS and then into pE100 (*YFP-Mtr-LTP*). After 48h of *A. tumefaciens*-mediated infiltration of *Mtr-WRKY-YFP-HA* or *YFP-Mtr-LTP* alone or together with *amiR-Rir-2216*, the silencing efficacy was evaluated by Western blot using an anti-YFP antibody.

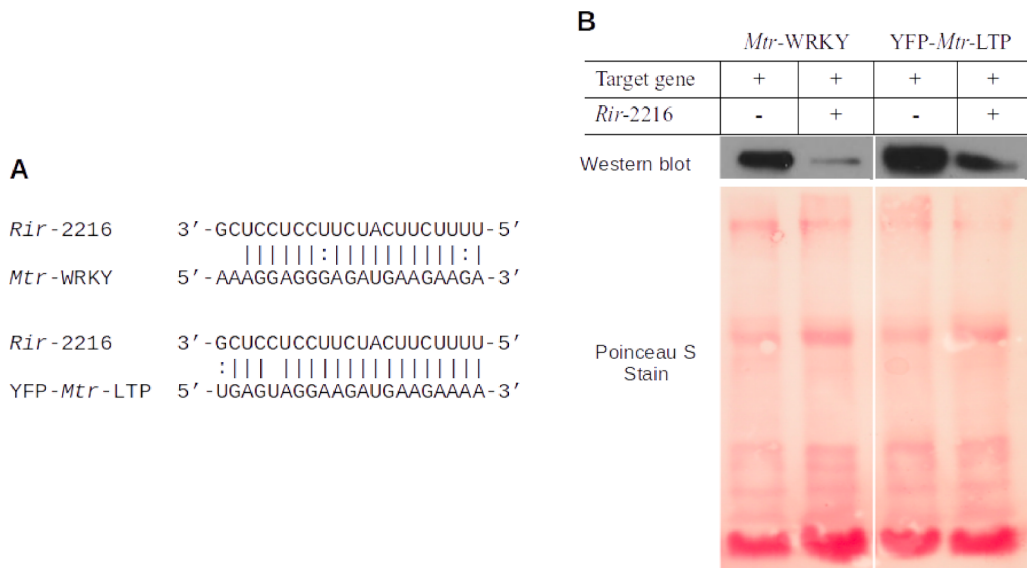


Figure 3. *Rir-2216* silences two of the *in silico* predicted *M. truncatula* target transcripts as revealed by co-expression assays in *N. benthamiana* leaves. **A)** The complementarity between *Rir-2216* with the two plant targets, *Mtr-WRKY* (AES66958) and *Mtr-LTP* (AES87505) is near perfect (Expectation = 1); **B)** Western blot of *Mtr-WRKY* (YFP-and HA- tagged; left) and of YFP tagged with the 21-nt long predicted target sequence of *Mtr-LTP* (right) expressed alone or together with *amiR-Rir-2216* (artificial miRNA containing *Rir-2216*). Ponceau S staining indicates that similar amounts of total proteins have been loaded in the different samples.

The YFP fusion proteins were clearly detected in the leaves infiltrated with the *Mtr-WRKY-YFP-HA* and *YFP-Mtr-LTP* *Agrobacterium* expressing cells; a strong reduction of protein accumulation was observed in leaves where

Mtr-WRKY-YFP-HA or YFP-*Mtr*-LTP were co-expressed with amiR-*Rir*-2216 (Fig. 3B). This result demonstrates that, at least in the transient expression assays, *Rir*-2216 is indeed able to silence two of the predicted *M. truncatula* genes. If *Rir*-2216 is transferred into plant cells during root colonization by *R. irregularis*, it can likely act as a silencing effector towards plant host genes.

Root colonization by *R. irregularis* is associated with lower transcript levels of *Rir*-2216 targets in cortical cells

At first qRT-PCR assays were used to monitor the transcript abundance of *Mtr*-WRKY, *Mtr*-LTP and *Mtr*-MHM17 (“plant/MHM17-10 protein”; AES59238), which is another *Rir*-2216 predicted target (Suppl. File 1), using RNA extracted from whole root of control non-mycorrhizal plants (RC) and mycorrhizal plants inoculated with *R. irregularis* (RM). No statistically significant differences were observed for *Mtr*-WRKY and *Mtr*-LTP genes, while an increased transcript level was reported for *Mtr*-MHM17 in RM condition compared to RC (Fig. 4A). *Mtr*-MHM17 is a gene containing a domain (Interpro: DUF761) characterizing the jasmonate-responsive defense gene “pathogen-associated molecular patterns-induced protein A70” from *Arabidopsis thaliana* (Truman et al., 2007). Mycorrhizal roots are a heterogeneous environment composed by different cell types (epidermal, cortical, etc.) which can be not colonized or colonized by different fungal structures (intercellular hyphae or arbuscules). These cell types likely express a different set of genes to sustain cell specificity and/or particular symbiotic functions. Gene expression profiles carried out on whole mycorrhizal roots can therefore give an incomplete picture, since the regulation associated to specific cell types can be lost due to dilution effects. To highlight potential cell type-specific expression profiles, we carried out qRT-PCR assay from different cell types collected from mycorrhizal and non mycorrhizal roots by means of the laser microdissector (LMD). The LMD has been in the last decade a powerful tool to analyse plant as well as fungal gene expression pattern associated to specific cell types in the AMS (Balestrini et al., 2007; Gaude et al., 2011).

A qRT-PCR was performed to analyze the transcript abundance of *Mtr*-WRKY, *Mtr*-LTP and *Mtr*-MHM17 in three different root cortical cell populations: arbusculated cells (Arb), non-colonized cells from the vicinity of arbusculated ones (MnM) and control cells from non-mycorrhizal roots (C). The hypothesis

was that if *Rir-2216* is transferred into host cells during AMS, a reduction in transcript abundance of its target genes should be observed.

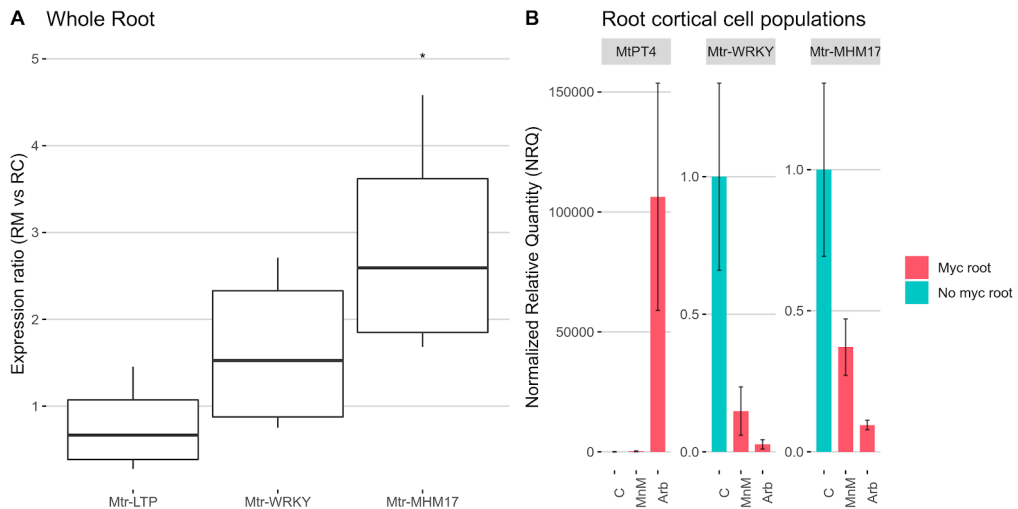


Figure 4. A) Whisker-box plot of the expression ratio (RM = “mycorrhizal root” vs RC = “control non-mycorrhizal root”) of *Mtr-LTP*, *Mtr-WRKY* and *Mtr-MHM17*. The asterisk highlights the gene (*Mtr-MHM17*) with significant differential expression level between the two conditions as calculated by REST2009. **B)** Relative abundance of *MtPT4*, *Mtr-WRKY* and *Mtr-MHM17* transcripts in 3 different populations of *M. truncatula* root cortical cells collected by means of LMD (C = control cells from non-mycorrhizal roots; MnM = non-colonized cells from the nearby of arbusculated ones; Arb = arbusculated cells); data are presented as average Normalized Relative Quantity (NRQ) \pm standard error.

As a first control step, we evaluated the expression level, in these three cell populations, of the plant phosphate transporter *MtPT4* which is a molecular marker of arbuscule-containing cells (Harrison et al., 2002; Javot et al., 2007). As expected, *MtPT4* was expressed only in arbusculated cells (Arb; Fig. 4B). We then proceeded to the qRT-PCR analysis for the three genes of interest: *Mtr-WRKY*, *Mtr-LTP* and *Mtr-MHM17*. Unfortunately, we could not detect *Mtr-LTP* transcripts in any of the root cell populations analyzed, probably due to its low abundance (data not shown). By contrast, we observed a decrease in *Mtr-WRKY* and *Mtr-MHM17* transcript abundances in MnM and Arb cell populations compared to C (Fig. 4B). This is in line with the hypothesis that *Rir-2216* could be transferred into host cells, where it could silence its host

target genes. Notably, the *Mtr-WRKY* and *Mtr-MHM17* transcript reductions were more pronounced in the Arb cell population (where the AMF shows a deep intimate contact with plant cells) than in the MnM one (Fig. 4B).

Based on these data, we can speculate that *Rir-2216* molecules are mainly transferred from AMF to plant into arbusculated cells where they trigger the silencing of *Mt-WRKY*. Once entered the root symplast, fungal sRNAs may then also spread into neighboring non-colonized cells where they also down-regulate its plant targets. In light also of *Rir-2216* low expression level (9 RPM), it is likely that the gene expression regulatory activity of *Rir-2216* towards host targets is limited to a low number of root cells (those directly in contact, or in the nearby, of fungal structures).

Generation of *M. truncatula* hairy roots expressing 3xHA-tagged AGO1

Further experiments are needed in order to validate this model. In particular, we should prove that the down-regulation of *Mtr-WRKY* and *Mtr-MHM17* in cortical cells from mycorrhizal roots is driven by the fungal *Rir-2216* and it is not just a mere consequence of endogenous host response to AMF colonization. For this purpose we therefore planned to perform immunoprecipitation experiments of *M. truncatula* AGOs with the aim to verify whether from mycorrhizal roots, it is possible to isolate fungal sRNAs (and specifically *Rir-2216*) bound to plant AGOs (Weiber et al., 2013). We started with the cloning of the cDNA sequence of *M. truncatula* AGO1 (*MtAGO1* tagged with a 3xHA at its 5'-end; 3xHA-*Mtr*-AGO1) under the 35S promoter into the binary vector PK7WG2D which is characterized, as selection marker, by a second expression cassette encoding for a GFP.

We performed an *Agrobacterium rhizogenes*-mediated transformation of *M. truncatula* and, so far, we obtained one stable line of 3xHA-*Mtr*-AGO1-expressing hairy roots. As a next step, these hairy roots will be colonized with *R. irregularis*. Then, immunoprecipitation of 3xHA-*Mtr*-AGO1, by means of anti-HA antibodies, will be performed from mycorrhizal and non-mycorrhizal hairy roots; the sequencing of the sRNAs present into this immunoprecipitate will allow us to demonstrate whether fungal sRNAs move from AMF to the host plant in AMS and bind plant AGO1.

M. truncatula genome contains 12 AGO encoding genes (Capitão et al., 2011): we decided to start this experiment with AGO1 since 1) in *A. thaliana* AGO1 is involved in PTGS by binding 21-nt long sequences (Axtell, 2013), that is the

same size of *Rir-2216*, and 2) AGO1 is the only member of the AGO family known so far to be involved in cross-kingdom RNAi (Weiberg et al., 2013; Wang et al., 2016; Cui et al., 2019; Ren et al., 2019).

3.3. Conclusions and Perspectives

The results presented in this work are consistent with the hypothetical movement of functional sRNAs from fungi to plants in AMS. In particular, we demonstrated that the *R. irregularis*-derived sRNA *Rir-2216* is able to silence two of its predicted plant transcripts (*Mtr-WRKY* and *Mtr-LTP*) in a transient co-expression assay. Interestingly, we also reported that two of the *Rir-2216* plant targets (*Mtr-WRKY* and *Mtr-MHM17*) show a reduced transcript accumulation in two populations of cortical cells from mycorrhizal roots compared to a counterpart from non-mycorrhizal roots. We finally proposed a model in which *Rir-2216* is transferred to host plant through the arbuscules, from where then it migrates to the near non-colonized host cells.

Unfortunately, none of the putative plant target genes analysed so far, *Mtr-WRKY*, *Mtr-LTP* and *Mtr-MHM17*, has been characterized in *M. truncatula*. However, at least on the basis of the literature data on homolog sequences studied in other species, they seem to be somehow involved in plant response to pathogens (Truman et al., 2007; Maldonado et al., 2002; Hahlbrock et al., 2003). It can be hypothesized that the down-regulation triggered by the fungal sRNA is important for the colonization process and/or the correct development of arbuscules. It is also worth noting that the expression of a LTP gene was shown to be regulated in rice roots in response to colonization by an AMF: transcript levels increased when the fungus formed hyphopodia and penetrates the root epidermis while decreased at the onset of the intercellular colonization of the root cortex (Blilou et al., 2000). Further investigations are needed to test these hypotheses.

As future experiments we therefore plan to generate *M. truncatula* plants constitutively expressing *Rir-2216* to evaluate (1) whether the predicted target genes of *Rir-2216* are actually down-regulated and, in this case, (2) to characterize the mycorrhizal phenotype of these transgenic plants. In the same line, it would be interesting also to evaluate the mycorrhizal phenotype of plants where the putative *Rir-2216* plant target genes (*Mtr-WRKY*, *Mtr-LTP* or *Mtr-MHM17*) are overexpressed or silenced.

3.4. Materials and Methods

RT-PCR on sRNAs

A portion of total RNA previously extracted (Silvestri et al., 2019) was used to reverse-transcribe and amplify sRNAs, following a end-point RT-PCR protocol (Varkonyi-Gasic et al., 2007). Briefly, each sRNA was reverse-transcribe using specific stem-loop oligonucleotides overlapping, at their 3'-ends, the last 6 nucleotides of the 3'-ends of sRNAs. The RT products were then amplified by means of PCR with a universal reverse primer (annealing on the conserved region of stem-loop sequence used for the reverse-transcription) and a sRNA-specific forward primer. All PCR runs were performed for 28-30 cycles and with an annealing-extension temperature of 60-66.1°C, depending on which sequence was analyzed. PCR reactions were analyzed by electrophoresis on 4% agarose gels. The list of used oligonucleotides is shown in Tab. 1.

Table 1. List of oligonucleotides used for RT-PCR on sRNAs

Name	Oligonucleotide
Universal-Rev-PCR	GTGCAGGGTCCGAGGT
Rir-2216-For-PCR	cggcggTTTTCTTCATCTTCC
Rir-2216-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACcggagga
Mtr-399a-For-PCR	cggcggUGCCAAAGGAGAUUU
Mtr-399a-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACcctgggc
Rir-196-For-PCR	cgcgCCCGAGGCTGGAAACAG
Rir-196-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACcctgct
Rir-434-For-PCR	cggcggTAATGATGAGACTTG
Rir-434-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACcctcca

Identification of *M. truncatula* target genes for *R. irregularis* sRNAs

We used only the most abundant sRNA from each *R. irregularis* sRNA-generating locus up-regulated in the intra-radical (RM) compared to the extra-radical (ERM) conditions, as analyzed in Silvestri et al. (2019), to predict targets in *M. truncatula* A17 transcriptome (v.4.0 cDNAs on EnsemblPlants database; Zerbino et al., 2018) through psRNAtarget (2017 update; Dai et al., 2018) with default parameters. We kept only the predictions with expectation lower than 2.5.

Co-expression assays

Following the procedure presented in Wang et al. (2017), the plant artificial miRNA vector containing *Rir-2216* (amiR-*Rir-2216*) was designed and generated according to the website WMD3 (<http://wmd3.weigelworld.org/cgi-bin/webapp.cgi>). The plasmid RS300 was used as template to amplify three fragments with primer pairs Oligo A and *Rir-2216* I (a), *Rir-2216* II and *Rir-2216* III (b), *Rir-2216* IV and Oligo B (c). The a, b, c fragments were used as templates primers to amplify fragment d (Oligo A and Oligo B), which exchanged *Rir-2216* into At-miR319a backbone. Fragment d was cloned into pENTR/SD-D-TOPO (Invitrogen), and next into the destination vector pEG100 by LR reaction (Invitrogen). *Mtr-WRKY* cDNA fragment, amplified with *Mtr-WRKY-For* (characterized by a 5'-cacc tag) and *Mtr-WRKY-Rev* oligonucleotides, was cloned into pENTR/SD-D-TOPO, and finally into the destination vector pEG101 (HA and YFP tags) by LR reaction (*Mtr-WRKY-YFP-HA*). The *Mtr-LTP* 21-nt long target sequence (5'-TGAGTAGGAAGATGAAGAAA-3') was cloned at the 3'-end of YFP fragment by YFP-For and *Mtr-LTP-YFP-Rev* oligonucleotides, then into pENTR/SD-D-TOPO and finally into pEG100 by LR reaction (YFP-*Mtr-LTP*). Oligonucleotide sequences are listed in Table 2. *A. tumefaciens* carrying amiR-*Rir-2216* was co-infiltrated with *Mtr-WRKY-YFP-HA* or YFP-*Mtr-LTP* into leaves of four-week-old *N. benthamiana* plants. Two days after infiltration, leaf samples were collected for Western blot analysis using anti-YFP antibodies.

Table 2. List of oligonucleotides used for co-expression assays

Name	Oligonucleotide
Oligo A	CTGCAAGGCGATTAAGTTGGGTAAC
Oligo B	caccGCGGATAACAATTTACACAGGAAACAG
<i>Rir-2216</i> I	gaTTTTCTTCATCTTCCTCCTCGtctctctttgtattcc
<i>Rir-2216</i> II	gaCGAGGAGGAAGATGAAGAAAAtcaaagagaatcaatga
<i>Rir-2216</i> -III	gaCGCGGAGGAAGATCAAGAAATtcacaggtcgtgatatg
<i>Rir-2216</i> -IV	gaATTTCTTGATCTTCCTCCGCGtctacatatattcct
WRKY-For	caccATGTACAAACGTAGATTCAACAC
WRKY-Rev	TCCTGTAATGCCGCTATACT
YFP-LTP-Rev	ttaTCTTTTCTTCATCTTCCTACTCATCT TCTGTACAGCTCGTCCATGC
YFP-For	caccATGGTGAGCAAGGGCGAGGAG

Quantitative RT-PCR in whole root samples

The same cDNA samples, from *M. truncatula* RM and RC roots, obtained in a previous study were used (Silvestri et al., 2019). qRT-PCR experiments were carried out in a final volume of 15 μ l containing 7.5 μ l of iTaq™ Universal SYBR Green Super- mix (Bio-Rad), 5.5 μ l of 0.8M primer mix and 2 μ l of 1:10 diluted cDNA. Amplification were run in a Rotor-Gene Q apparatus (Qiagen) using the following program: 5 min pre-incubation at 95 °C and 40 cycles of 30 s at 95 °C, 30 s at 60–64 °C. Each amplification was followed by melting curve analysis (60–94 °C) with a heating rate of 0.5 °C every 15 s. We used three biological replicates each with two technical replicates. Relative expression and statistical analyses were performed by REST2009, using as reference gene the plant *MtTEF*. The list of oligonucleotides is given in Table 3.

Table 3. List of oligonucleotides used for qRT-PCR in whole root samples

Name	Oligonucleotide
Mtr-WRKY-real-For2	GGGAGATGAAGAAGAGGGTGG
Mtr-WRKY-real-Rev2	TCTCCAAGCCCATGAATCCG
Mtr-LTP-For4	ACGCCACATTGTTGTGAAGC
Mt-LTP-Rev4	ACATAAGGCAAAGAGTGCAGC
Mtr-MHM17-real-For	TCTCTCCAATCTCCTCCACCC
Mtr-MHM17-real-Rev	AGCTGTAGAGACCAGCCCAG
MtTEF-real-For	AAGCTAGGAGGTATTGACAAG
MtTEF-real-Rev	ACTGTGCAGTAGTACTTGGTG
MtPT4-real-For	TCGCGCGCCATGTTTGTGT
Mt-PT4-real-Rev	CGCAAGAAGAAATGTTAGCCC

Quantitative RT-PCR on laser microdissected cells

As described in Belmondo et al. (2014), *M. truncatula* roots were colonized by *R. irregularis* with the millipore sandwich system (Giovannetti et al., 1993), cut into 5–10 mm-long pieces, treated with ethanol-glacial acetic acid (3:1) under vacuum for 30min and placed at 4°C overnight. Roots were subsequently dehydrated in a graded series of ethanol (50%–70%–90% in sterilized water and 100% twice) followed by Neoclear (twice) with each step on ice for 30 min; neoclear was gradually replaced with paraffin. A LeicaAS LMD system (Leica Microsystem) was used to collect cortical cells from paraffin root sections as described by Balestrini et al. (2007). Three cell populations were assayed: arbusculated cells (Arb), non-colonized cells from the vicinity of

arbusculated ones (MnM) and control cells from non-mycorrhizal roots (C). Two different biological replicates of 1500–2000 microdissected cells for each condition were collected and processed for RNA extraction. RNA was extracted from dissected cells using the PicoPure kit protocol (Arcturus Engineering). A DNase treatment was performed using an RNA-free DNase Set (Qiagen) in a Pico Pure column, according to the manufacturer's instructions and RNAs were eluted in 21µl of sterile water. DNA contaminations were assessed using the MtTEF primers described above (Tab. 3). All qRT-PCR assays were carried out using the One Step qRT-PCR kit (Qiagen). Reactions with specific primers were carried out in a final volume of 10 µl containing 2 µl of 5×buffer, 0.4 µl of 10 mM dNTPs, 1µl of each primer 10 mM, 0.2 µl of One Step RT-PCR enzyme mix, and 1 µl of a total RNA diluted 1:2-1:5, depending on the sample. The samples were incubated for 30 min at 50°C, followed by 15 min incubation at 95°C. Amplification reactions were run for 40 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 40 s. Normalized Relative Quantity (NRQ) for each transcript (*MtPT4*, *Mtr-WRKY* and *Mtr-MHM17*) in each condition (Arb, MnM, C) was calculated on an Excel spreadsheet using as reference MtTEF. The list of oligonucleotides is given in Table 3.

Generation of *M. truncatula* hairy roots carrying 3xHA-Mtr-AGO1

M. truncatula cDNA was used as template to amplify two overlapping fragments, with primer pairs *Mtr*-HA-AGO1-For and *Mtr*-AGO1-Rev-mid (a) and *Mtr*-AGO1-For-mid and *Mtr*-AGO1-Rev (b), that were used as template to amplify the full *MtrAGO1* fragment (carrying a 3xHA tag at 5'-end; 3xHA-Mtr-AGO1) with the primer pair *Mtr*-HA-AGO1-For and *Mtr*-AGO1-Rev. The 3xHA-Mtr-AGO1 was cloned into pENTR/SD-D-TOPO and then into the destination vector PK7WG2D by LR reaction. *A. rhizogenes* carrying the final vector was used to generate *M. truncatula* hairy roots as described in Chabaud et al. (2006). The list of oligonucleotides is given in Table 4.

Table 4. List of oligonucleotides used for 3xHA-Mtr-AGO1 cloning

Name	Oligonucleotide
Mtr-HA-AGO1-For	CACCatgTATCCTTATGATGTACCTGATTATGCCTACCCATACGACGTTCCAGACTACGCTTACCCATACGACGTTCCAGACTACGCTgtcaggaagaggagaactgatg
Mtr-AGO1-Rev	TCAACAGTAGAACATGACCTTC
Mtr-AGO1-For-mid	GTCACACATCGTGGCAATATG

Mtr-AGO1-Rev-mid	GACCAGCAATGCGATATTTTC
------------------	-----------------------

3.5. References

- Axtell, M. J. (2013). Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol.* 64, 137–159. doi:10.1146/annurev-arplant-050312-120043.
- Balestrini, R., Gómez-Ariza, J., Lanfranco, L., and Bonfante, P. (2007). Laser microdissection reveals that transcripts for five plant and one fungal phosphate transporter genes are contemporaneously present in arbusculated cells. *Mol. Plant-Microbe Interact.* 20, 1055–1062. doi:10.1094/MPMI-20-9-1055.
- Blilou, I., Ocampo, J. A., and García-Garrido, J. M. (2000). Induction of Ltp (lipid transfer protein) and Pal (phenylalanine ammonia-lyase) gene expression in rice roots colonized by the arbuscular mycorrhizal fungus *Glomus mosseae*. *J. Exp. Bot.* 51, 1969–1977. doi:10.1093/jexbot/51.353.1969.
- Branscheid, A., Sieh, D., Datt Pant, B., May, P., Devers, E. A., Elkrog, A., et al. (2010). Expression pattern suggests a role of MiR399 in the regulation of the cellular response to local Pi increase during arbuscular mycorrhizal symbiosis. *Mol. Plant-Microbe Interact.* 23, 915–926. doi:10.1094/MPMI-23-7-0915.
- Capitão, C., Paiva, J. A. P., Santos, D. M., and Fevereiro, P. (2011). In *Medicago truncatula*, water deficit modulates the transcript accumulation of components of small RNA pathways. *BMC Plant Biol.* 11, 79. doi:10.1186/1471-2229-11-79.
- Chiou, T., Aung, K., Lin, S., Wu, C., Chiang, S., and Su, C. (2006). Regulation of Phosphate Homeostasis by MicroRNA in *Arabidopsis*. *Plant Cell* 18, 412–421. doi:10.1105/tpc.105.038943.1.
- Cui, C., Wang, Y., Liu, J., Zhao, J., Sun, P., and Wang, S. (2019). A fungal pathogen deploys a small silencing RNA that attenuates mosquito immunity and facilitates infection. *Nat. Commun.* 10, 4298. doi:10.1038/s41467-019-12323-1.
- Dai, X., Zhuang, Z., and Zhao, P. X. (2018). PsRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* 46, W49–W54. doi:10.1093/nar/gky316.
- Gaude, N., Bortfeld, S., Duensing, N., Lohse, M., and Krajinski, F. (2011). Arbuscule-containing and non-colonized cortical cells of mycorrhizal roots undergo extensive and specific reprogramming during arbuscular mycorrhizal development. *Plant J.* 69, 510–528. doi:10.1111/j.1365-313X.2011.04810.x.
- Hahlbrock, K., Heise, A., Liedgens, H., Bednarek, P., Ciolkowski, I., Schmelzer, E., et al. (2003). Non-self recognition, transcriptional reprogramming, and secondary metabolite accumulation during plant/pathogen interactions. *Proc. Natl. Acad. Sci. U. S. A.* 100, 14569–14576.
- Harrison, M. J., Dewbre, G. R., and Liu, J. (2002). A phosphate transporter from *Medicago truncatula* involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. *Plant Cell* 14, 2413–29. doi:10.1105/TPC.004861.

- Helber, N., Wippel, K., Sauer, N., Schaarschmidt, S., Hause, B., and Requena, N. (2011). A versatile monosaccharide transporter that operates in the arbuscular mycorrhizal fungus *Glomus* sp is crucial for the symbiotic relationship with plants. *Plant Cell* 23, 3812–3823. doi:10.1105/tpc.111.089813.
- Hogekamp, C., Arndt, D., Pereira, P. A., Becker, J. D., Hohnjec, N., and Kuster, H. (2011). Laser microdissection unravels cell-type-specific transcription in arbuscular mycorrhizal roots, including CAAT-Box transcription factor gene expression correlating with fungal contact and spread. *Plant Physiol.* 157, 2023–2043. doi:10.1104/pp.111.186635.
- Huang, C.-Y., Wang, H., Hu, P., Hamby, R., and Jin, H. (2019). Small RNAs – big players in plant-microbe interactions. *Cell Host Microbe* 26, 173–182. doi:10.1016/j.chom.2019.07.021.
- Ivanov, S., Austin, J., Berg, R. H., and Harrison, M. J. (2019). Extensive membrane systems at the host–arbuscular mycorrhizal fungus interface. *Nat. Plants* 5, 194–203. doi:10.1038/s41477-019-0364-5.
- Javot, H., Penmetsa, R. V., Terzaghi, N., Cook, D. R., and Harrison, M. J. (2007). A *Medicago truncatula* phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1720–1725. doi:10.1073/pnas.0608136104.
- Kikuchi, Y., Hijikata, N., Ohtomo, R., Handa, Y., Kawaguchi, M., Saito, K., et al. (2016). Aquaporin-mediated long-distance polyphosphate translocation directed towards the host in arbuscular mycorrhizal symbiosis: application of virus-induced gene silencing. *New Phytol.* 211, 1202–1208. doi:10.1111/nph.14016.
- Maldonado, A. M., Doerner, P., Dixon, R. A., Lamb, C. J., and Cameron, R. K. (2002). A putative lipid transfer protein involved in systemic resistance signalling in *Arabidopsis*. *Nature* 419, 399–403. doi:10.1038/nature00962.
- Ren, B., Wang, X., Duan, J., and Ma, J. (2019). Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation. *Science*. 365, 919–922. doi:10.1126/science.aav8907.
- Roth, R., Hillmer, S., Funaya, C., Chiapello, M., Schumacher, K., Lo Presti, L., et al. (2019). Arbuscular cell invasion coincides with extracellular vesicles and membrane tubules. *Nat. Plants* 5, 204–211. doi:10.1038/s41477-019-0365-4.
- Silvestri, A., Fiorilli, V., Miozzi, L., Accotto, G. P., Turina, M., and Lanfranco, L. (2019). In silico analysis of fungal small RNA accumulation reveals putative plant mRNA targets in the symbiosis between an arbuscular mycorrhizal fungus and its host plant. *BMC Genomics* 20, 169. doi:10.1186/s12864-019-5561-0.
- Truman, W., Bennet, M. H., Kubigsteltig, I., Turnbull, C., and Grant, M. (2007). *Arabidopsis* systemic immunity uses conserved defense signaling pathways and is mediated by jasmonates. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1075–1080. doi:10.1073/pnas.0605423104.
- Tsuzuki, S., Handa, Y., Takeda, N., and Kawaguchi, M. (2016). Strigolactone-induced putative secreted protein 1 is required for the establishment of symbiosis by the arbuscular mycorrhizal fungus *Rhizophagus irregularis*. *Mol. Plant-Microbe Interact.* 29, 1–59. doi:10.1094/MPMI-10-15-0234-R.

-
- Varkonyi-Gasic, E., Wu, R., Wood, M., Walton, E. F., and Hellens, R. P. (2007). Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods* 3, 12. doi:10.1186/1746-4811-3-12.
- Voß, S., Betz, R., Heidt, S., Corradi, N., and Requena, N. (2018). RiCRN1, a crinkler effector from the arbuscular mycorrhizal fungus *rhizophagus irregularis*, functions in arbuscule development. *Front. Microbiol.* 9, 1–18. doi:10.3389/fmicb.2018.02068.
- Wang, M., Weiberg, A., Dellota, E., Yamane, D., and Jin, H. (2017). *Botrytis* small RNA Bc-siR37 suppresses plant defense genes by cross-kingdom RNAi. *RNA Biol.* 14, 421–428. doi:10.1080/15476286.2017.1291112.
- Wang, M., Weiberg, A., Lin, F. M., Thomma, B. P. H. J., Huang, H. Da, and Jin, H. (2016). Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. *Nat. Plants* 2, 16151. doi:10.1038/nplants.2016.151.
- Weiberg, A., Wang, M., Lin, F. M., Zhao, H., Zhang, Z., Kaloshian, I., et al. (2013). Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science*. 342, 118–123. doi:10.1126/science.1239705.
- Xie, X., Lin, H., Peng, X., Xu, C., Sun, Z., Jiang, K., et al. (2016). Arbuscular mycorrhizal symbiosis requires a phosphate transceptor in the *Gigaspora margarita* fungal symbiont. *Mol. Plant* 9, 1583–1608. doi:10.1016/j.molp.2016.08.011.
- Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic Acids Res.* 46, D754–D761. doi:10.1093/nar/gkx1098.

R. irregularis sRNA Rir-2216 is likely involved in cross-kingdom RNAi

CHAPTER 4

4. The virome of the arbuscular mycorrhizal fungus *Gigaspora margarita* reveals the first report of DNA fragments corresponding to replicating non-retroviral RNA viruses in fungi

Massimo Turina¹, Stefano Ghignone¹, Nausicaa Astolfi¹, Alessandro Silvestri², Paola Bonfante² and Luisa Lanfranco²





¹Institute for Sustainable Plant Protection – CNR Torino, Strada delle Cacce 73, 10131 Torino, Italy

²Department of Life Sciences and Systems Biology, University of Torino, Viale P.A. Mattioli 25, 10125 Torino, Italy

Published on *Environmental Microbiology*

Turina et al. Environ. Microbiol. (2018) doi:10.1111/1462-2920.14060.

The virome of the arbuscular mycorrhizal fungus *Gigaspora margarita* reveals the first report of DNA fragments corresponding to replicating non-retroviral RNA viruses in fungi

Massimo Turina ¹, Stefano Ghignone ¹,
Nausicaa Astolfi,² Alessandro Silvestri,²
Paola Bonfante ² and Luisa Lanfranco ^{2*}

¹Institute for Sustainable Plant Protection – CNR Torino,
Strada delle Cacce 73, 10131 Torino, Italy.

²Department of Life Sciences and Systems Biology,
University of Torino, Viale P.A. Mattioli 25, 10125 Torino,
Italy.

Summary

Arbuscular Mycorrhizal Fungi (AMF) are key components of the plant microbiota. AMF genetic complexity is increased by the presence of endobacteria, which live inside many species. A further component of such complexity is the virome associated to AMF, whose knowledge is still very limited. Here, by exploiting transcriptomic data we describe the virome of *Gigaspora margarita*. A BLAST search for viral RNA-dependent RNA polymerases sequences allowed the identification of four mitoviruses, one Ourmia-like narnavirus, one Giardia-like virus, and two sequences related to *Fusarium graminearum* mycoviruses. Northern blot and RT-PCR confirmed the authenticity of all the sequences with the exception of the *F. graminearum*-related ones. All the mitoviruses are replicative and functional since both positive strand and negative strand RNA are present. The abundance of the viral RNA molecules is not regulated by the presence or absence of *Candidatus Glomeribacter gigasporarum*, the endobacterium hosted by *G. margarita*, with the exception of the Ourmia-like sequence which is absent in bacteriucured spores. In addition, we report, for the first time, DNA fragments corresponding to mitovirus sequences associated to the presence of viral RNA. These sequences are not integrated in the mitochondrial

DNA and preliminary evidence seems to exclude integration in the nuclear genome.

Introduction

The majority of land plants, including many crops and horticultural species, establish root symbiotic interactions with a small group of soil fungi, the so called arbuscular mycorrhizal fungi (AMF), which belong to the Glomeromycotina (Spatafora *et al.*, 2016). As very ancient and widespread in nature, AMF are key components of the plant microbiota. The benefits that AMF provide to host plants, which include an improved mineral nutrition and an increased tolerance to biotic and abiotic stresses, have raised the interest towards their exploitation as key components of sustainable low input agricultural practices (Rodriguez and Sanders, 2015; Berruti *et al.*, 2015).

The AMF obligate biotrophic nature and the lack of stable genetic transformation protocols have made them a recalcitrant biological system to study. The recent publications of genomic data for *Rhizophagus irregularis* (Tisserant *et al.*, 2013; Lin *et al.*, 2014; Ropars *et al.*, 2016) and gene repertoires of other species such as *R. clarus* (Sędziewska Toro and Brachman, 2016), *Gigaspora margarita* (Salvioli *et al.*, 2016) and *G. rosea* (Tang *et al.*, 2016) have shed light on the biology and evolution of AMF (Kamel *et al.*, 2017). These findings have allowed to propose a new phylogenetic classification (Spatafora *et al.*, 2016) and to infer the occurrence of a sexual cycle in AMF (Corradi and Brachmann, 2017), for long considered clonal organisms. *R. irregularis* possesses a large genome, compared to other fungi, with about 28 000 protein encoding genes and rich in transposable elements (Tisserant *et al.*, 2013; Lin *et al.*, 2014). One striking feature is the almost complete lack of glycoside hydrolases, a sign of the inability of the fungus to degrade plant cell wall polysaccharides, possibly a strategy to guarantee an intimate and compatible interaction.

The genetic complexity of AMF is also increased by the presence of endobacteria living inside hyphae and spores

Received 28 November, 2017; revised 23 January, 2018; accepted 28 January, 2018. *For correspondence. E-mail: luisa.lanfranco@unito.it; Tel. 00390116705969; Fax 00390116705962.

of many AMF (Bonfante and Desirò, 2017). Two types of endobacteria have been described in AMF: rod-shaped and Gram-negatives, associated to members of the Gigasporaceae family (Ghignone *et al.*, 2012) and coccoid Mollicutes-related endobacteria, distributed across different lineages of AM fungi (Naumann *et al.*, 2010; Desirò *et al.*, 2014; Torres-Cortés *et al.*, 2015). To date, only *Candidatus Glomeribacter gigasporarum*, the rod-shaped bacterium hosted by *Gigaspora margarita* has been investigated in more detail. Thanks to the comparison with a cured, endobacteria-free, strain (Lumini *et al.*, 2007) it has been demonstrated that the endobacterium enhances fungal sporulation, bioenergetic capacity by priming mitochondrial metabolic pathways and ability to detoxify reactive oxygen species possibly leading to an improved ecological fitness of the fungus (Salvioli *et al.*, 2016; Vannini *et al.*, 2016).

All major fungal lineages were shown to possess additional genetic components among which are viruses also called mycoviruses (Ghabrial *et al.*, 2015; Son *et al.*, 2015). In most cases the genome of mycoviruses is an RNA molecule, mainly a double-stranded RNA (dsRNA), containing at least a gene encoding an RNA-dependent RNA polymerase (RdRp), which is necessary for replication (Ghabrial and Suzuki, 2009). Mycoviruses are transmitted horizontally via hyphal anastomosis between vegetatively compatible individuals or transmitted vertically through asexual and, less frequently, through sexual spores (Hillman *et al.*, 2004; Ghabrial *et al.*, 2015). External route of infection is apparently lacking; however, recent studies suggest that mycophagous insects function as mycovirus transmission vectors in the case of ssDNA mycoviruses (Liu *et al.*, 2016). The infection by mycoviruses is often asymptomatic, but in some specific conditions alterations in the phenotype of the host fungus can be observed (van Diepeningen *et al.*, 2006; Nerva *et al.*, 2017). In some cases higher order biological interactions are also influenced; for example, the presence of a virus can lead to attenuation of virulence in a plant pathogenic fungus (Choi and Nuss, 1992) or can provide the capability to enhance thermal tolerance of the host plant to an endophytic fungus (Márquez *et al.*, 2007).

The knowledge of mycoviruses in mycorrhizal fungi is still very limited. In the ectomycorrhizal fungus *Tuber aestivum*, three mycoviruses - a mitovirus (Stielow *et al.*, 2011a), a totivirus (Stielow and Menzel, 2010) and an endornavirus (Stielow *et al.*, 2011b) - have been described but their biological role was not characterized. Concerning AMF, Ikeda and colleagues (2012) demonstrated, for the first time, the presence of mycovirus-related dsRNAs, in the mycelium of *Glomus sp.* (now named *Rhizophagus*) strain RF1. A 4.557 nucleotides segment, called GRF1V-M, which could not be phylogenetically assigned to known genera of mycovirus, was characterized in detail.

The GRF1V-M encodes an RdRp and a protein of unknown function. Remarkably, by subculturing single spores, a fungal virus-free line was obtained: the absence of the virus led to the production of a higher number of spores and to an enhanced stimulation of plant growth compared to the GRF1V-M-positive line (Ikeda *et al.*, 2012). In a second work the same group has characterized by deep sequencing of dsRNA in the AMF *Rhizophagus clarus* strain RF1, a sequence showing similarity to RdRp of mitoviruses (Kitahara *et al.*, 2014). In this case, a biological function has not been reported.

Here, by exploiting recently published transcriptomic data (Salvioli *et al.*, 2016) we describe for the first time the components of the virome of the AMF *G. margarita*, which is phylogenetically distantly related to *Rhizophagus* species. We provide *in silico* and experimental evidence of the existence of a population of six viral sequences (4 mitoviruses, one Giardia-like and one Ourmia-like virus). With one exception (the Ourmia-like virus), we found that the abundance of the viral RNA molecules was not regulated by the presence of the endobacterium hosted by *G. margarita*. Interestingly, we for the first time report DNA fragments corresponding to mitovirus sequences associated to the presence of their corresponding viral RNA: these sequences are not integrated in the mitochondrial or nuclear genome, but likely exist as extrachromosomal fragments.

Results

Identification of viral sequences from *Gigaspora margarita* transcriptome

Contigs assembled from RNAseq data for the AMF *Gigaspora margarita* BEG34 (Salvioli *et al.*, 2016) were searched for sequences annotated as RdRp; these are proteins performing enzymatic activities typically encoded by RNA viral genomes, easily distinguishable from host RdRp based on homology to other viral or host (fungal) RdRps. Eight sequences were retrieved and a BLAST search allowed us to identify four mitoviruses, one Ourmia-like virus, one Giardia-like dsRNA virus and two partial sequences related to *Fusarium graminearum* mycovirus 1 (FgV1) and *Fusarium graminearum* mycovirus (FgV3) (Supporting Information Table S1 and Fig. S1). The four putative mitoviral sequences and the Ourmia-like contig encode for a single ORF for each genome, which, based on the presence of the conserved GDD motif, can be annotated as putative RdRps (Table 1 and Supporting Information Fig. S1). In the case of the Giardia-like virus, genomic sequence predicts the presence of two ORFs: the first ORF encodes a putative protein of 418 aa with no similarity with proteins currently present in databases; the second ORF of 1022 aa shows similarity to RdRps of another Giardia-like virus isolated in the AMF *Rhizophagus*

Table 1. Molecular features of the viral sequences identified in the *G. margarita* transcriptome.

Virus	Contig length (nt)	ORF	Position of AUG	Stop codon	Protein length (aa)
33086 contig mitovirus	3243	RdRp	473	2855	794
34036 contig mitovirus	3394	RdRp	305	3263	986
34470 contig mitovirus	3387	RdRp	282	3312	1010
34875 contig mitovirus	3389	RdRp	276	3300	1008
36178 contig Ourmia-like	3144	RdRp	269	3062	931
33452 contig Giardia-like	4947	ORF1	311	1567	418
		RdRp	1611	4677	1022

sp. (Ikeda *et al.*, 2012) (Supporting information Table 1). Details of the genome organization of the viral sequences, which were then experimentally confirmed (see below), are given in Table 1.

To confirm the authenticity (presence as biological entities, i.e., RNA molecules, and not as mere *in silico* artifacts) of the sequences assembled from the transcriptome, RT-PCR assays were performed using RNA from germinating spores of the same *G. margarita* strain with specific primers (Supporting Information Table S2) designed to amplify a cDNA fragment in the size-range of about 500–700 bp. A fragment of the expected size was obtained for each of the four mitoviruses and the

Giardia-like virus (Fig. 1A). The sequencing of recombinant plasmids, after cloning the PCR fragments confirmed the identity of the sequence of the PCR product with that predicted *in silico* (data not shown). More primer pairs spanning different regions of the *in silico* predicted genomes were designed for the remaining sequences (those that could not be confirmed in the first round of PCR, such as contig 23972, 39980 and 36178x, the FgV1 and FgV3 and the Ourmia-like related sequences respectively) and tested again in RT-PCR assays. No amplified fragment was obtained for the 23972 and 39980 contigs, the two *F. graminearum* mycovirus-related sequences. Only for the Ourmia-like sequence (contig 36178) we could confirm by RT-PCR and sequencing the presence of an RNA molecule corresponding to a short portion (324 bp) identical to the virus RdRp predicted *in silico* (Fig. 1B). Although the 36178 contig obtained *in silico* (Ourmia-like putative virus) encodes a long ORF, the amplified segment corresponds only to conserved RdRp motifs, whereas the rest of the sequence does not show similarity to existing RdRps; several attempts to verify the presence of other cDNA fragments corresponding to other regions of the Ourmia-like contig extending outside the conserved GDD domain failed: we therefore considered that the *in silico* assembled sequence may be a chimera generated by an artifact of RNAseq assembly, and that only the conserved region we could amplify is corresponding to a specific Ourmia-like RNA molecule.

To confirm that each identified viral sequence encodes for an active RdRp, northern blot assays were carried out on *G. margarita* germinating spores using positive and negative strand riboprobes. Northern blot allowed to detect genomic (+sense) RNA corresponding to the four mitoviruses and the narnavirus-like sequence (Ourmia-like contig 36178) (Fig. 2), but we could provide evidence for active replication only for the mitoviruses since both full length positive strand and negative strand RNA were detected in RNA extracts; in addition, they seem rather more abundant (Fig. 2A–D) compared to the other virus tested. For the Ourmia-like sequence only the positive sense-detecting probe gave a hybridization signal; no signal was obtained for the negative sense-detecting probe

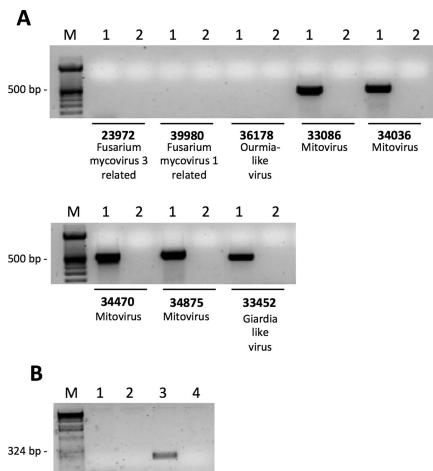


Fig. 1. A. Gel electrophoresis of RT-PCR products amplified from cDNA from *G. margarita* germinating spores (1) with the following primer pairs specific for the putative viral sequences: 33086genF1 and 33086genR1, 34036Fgen and 34036Rgen, 34770Fgen and 34770Rgen, 34875Fgen and 34875Rgen, 36178Fgen and 36178Rgen, 33452Fgen and 33452Rgen, 23972Fgen and 23972Rgen, 39980Fgen and 39980Rgen. M: 100 bp ladder (Invitrogen); no cDNA sample (2).

B. Gel electrophoresis of PCR products amplified from *G. margarita* DNA (1); *G. margarita* DNA treated with DNase (2); *G. margarita* cDNA from germinating spores (3); no DNA sample (4) with primers specific for the Ourmia-like sequence (36178 contig). M: 1 kb ladder (Invitrogen).

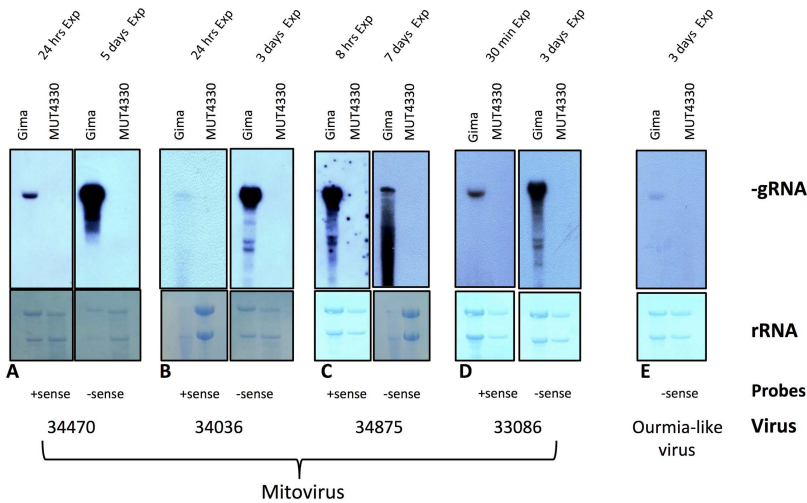


Fig. 2. Northern blot analysis of *G. margarita*-associated viral sequences. Time of autoradiography exposure (Exp) is indicated on top of each panel. MUT4330 (total RNA from an unrelated filamentous fungus) was used as negative control in the hybridization experiment. Lower panels are methylene blue stained membranes to show ribosomal RNA loading. gRNA, genomic RNA; rRNA, ribosomal RNA.

probably due to the low abundance of the target sequence (Fig. 2E), even after 20 days exposure (data not shown).

No hybridization signal was detected for the Giardia-like sequence with both probes, suggesting that the abundance of the RNA was below the threshold for northern detection in our experiments (data not shown). The very low abundance of RNA corresponding to the Ourmia-like virus and the Giardia-like virus was confirmed by qRT-PCR assays (see below).

Phylogenetic placement of viral sequences

We then further characterized our viral sequences with a phylogenetic analysis comparing conserved regions of the RdRp, with those present in the databases. We first aligned the *G. margarita* mitovirus and Ourmia-like virus to a data set of RdRp that comprises the family *Ourmiaviridae*, *Leviviridae*, *Narnaviridae* and a number of related viruses still not classified taxonomically that resulted from NGS virome characterization (Shi *et al.*, 2016). The conserved RdRp region among this wide data set is limited to 109 amino acids surrounding the GDD catalytic motif. This phylogenetic analysis allows to show that *G. margarita* putative mitoviruses are indeed members of the *Mitovirus* genus, whereas the Ourmia-like fragment present in *G. margarita* belongs to a clade of viruses still not classified taxonomically (Supporting Information Fig. S2), which we have previously proposed to form a new genus called Ourmia-like viruses (Turina *et al.*, 2017).

A further more defined phylogenetic analysis, that includes only mitoviruses, was carried out on aligned sequences spanning 467 amino acids using an endogenized plant mitovirus fragment from *Raphanus sativus* as outgroup (Fig. 3). This analysis showed that mitoviruses

infecting *G. margarita* are phylogenetically diverse. In particular, the 33086 is distantly related to the other three sequences (34036, 34470 and 34875) which instead form a statistically well supported clade. Such clade is separated from the other statistically supported mitovirus clades (Fig. 3). Indeed, contig 33086 shares the highest RdRp protein identity score with *Botrytis cinerea* mitovirus (38.0%), while the highest identity value shared with the other *G. margarita* mitoviruses is 30.1% (33086 vs. 34875). The other three mitoviruses, instead, share higher percentage identity among each other respectively from 32.0 (34036 vs. 34875) to 53.5 (34770 vs. 34036) (Table 2 and Supporting Information Table S3). Interestingly, in all the 4 mitovirus sequences 100% of tryptophan (W) residues are encoded by the TGG codon (Supporting Information Fig. S1): therefore, a functional RpRd can be hypothetically translated both in the cytosol and the mitochondria.

A phylogenetic tree for the Giardia-like viral sequence was also generated including the most closely related members of dsRNA virus families (*Totiviridae* and *Partitiviridae*); as expected, virus RdRps belonging to the *Totiviridae* and *Partitiviridae* are grouped in statistically well supported clades; the rest of the aligned sequences (including the *G. margarita* Giardia-like virus) forms a clade with a relatively low statistical support value (39%) in bootstrap analysis (Fig. 4).

Are the viral sequences endogenized into the *G. margarita* genome?

A widespread endogenization of mitoviral sequences in plant genomes was recently reported (Bruenn *et al.*, 2015) and this prompted us to look for possible genome

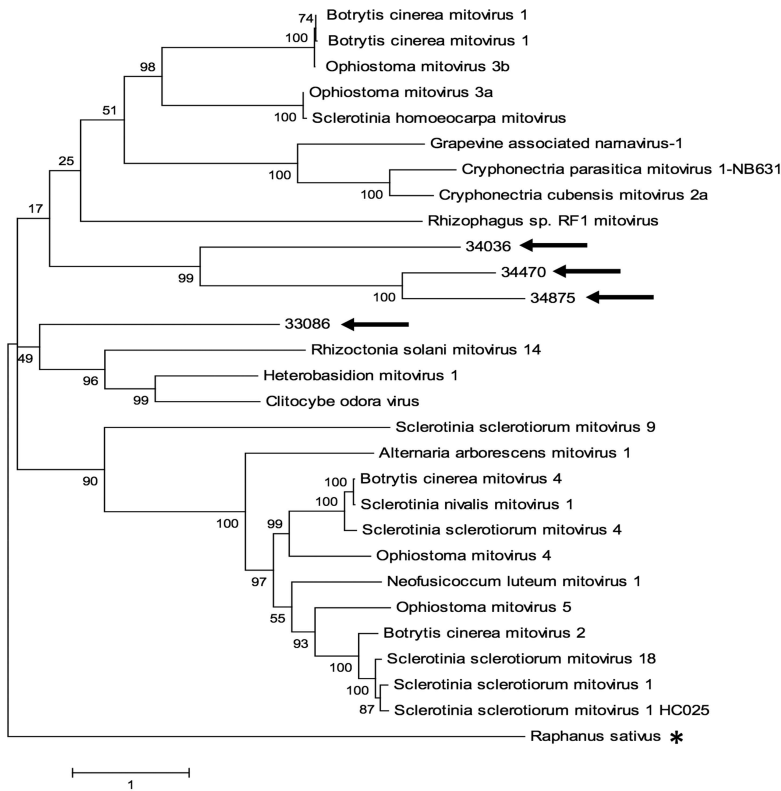


Fig. 3. Molecular phylogenetic analysis by Maximum Likelihood method of Mitovirus sequences (*G. margarita* sequences are indicated by arrows). The tree with the highest log likelihood (-22958.4232) is shown. Initial tree(s) for the heuristic search were obtained by applying the Neighbour-Joining method to a matrix of pairwise distances estimated using a JTT model. A discrete Gamma distribution was used to model evolutionary rate differences among sites: 5 categories (+G, parameter = 1.0199). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 3.8500% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All positions with less than 95% site coverage were eliminated. There were a total of 467 positions in the final data set. The symbol "*" refers to an endogenized fragment of a putative plant mitovirus in the genome of the plant *Raphanus sativus* which was used as outgroup.

endogenization events of the viral sequences we identified in *G. margarita*.

We performed PCR assays on DNA extracted from *G. margarita* spores with the same specific primers that were used to amplify the cDNA corresponding to the RNA (Supporting Information Table S1). In case of positive outcome, to confirm that the PCR products were indeed originated from DNA, a control sample using DNase-treated genomic DNA was also analysed in parallel to exclude any spurious activity of the DNA polymerase that might function as reverse transcriptase. A PCR product of

the expected size, which was absent in the DNase-treated sample, was obtained for the four mitoviruses and for the FgV3-related 23972 sequence (Fig. 5). These PCR products could also be amplified on RNase-treated DNA preparations (data not shown). Therefore, for these five virus-like sequences a phenomenon of genome endogenization can be hypothesized.

To further explore this issue we checked the sequence of the previously characterized *G. margarita* mitochondrial genome (BEG34 strain, the same used in this study) (Pelin *et al.*, 2012): no viral sequence was found within the complete mitochondrial DNA. We then tested the hypothesis that these DNA sequences corresponding to viral RdRps could be endogenized into the nuclear genome. A draft genome sequence of *G. margarita* (BEG 34 strain) is available (Ghignone, Venice, Salvioli, Bonfante, unpublished). We mapped the genome sequencing reads on viral contigs using BWA software. The DNA sequence encoding the elongation factor, used as a positive control, was well represented (590 paired end reads). We also had a confirmation of the genome endogenization for the 23972 sequence (FgV3-related) for which 164 paired end reads were found. Furthermore, BLASTN analysis of the partial

Table 2. Percentages similarity (italics) and identity (underlined) of pairwise alignments of the four *G. margarita* mitovirus sequences along the entire protein sequence.

Mitovirus contigs	33086	34036	34470	34875
33 086		<i>34%</i>	<i>32%</i>	<i>33%</i>
34 036	<i>21%</i>		<i>38%</i>	<i>40%</i>
34 470	<i>19%</i>	<i>23%</i>		<i>56%</i>
34 875	<i>20%</i>	<i>24%</i>	<i>39%</i>	

Identity and similarity among pairwise alignments were calculated using the Needleman-Wunsch Global Align Protein Sequence algorithm.

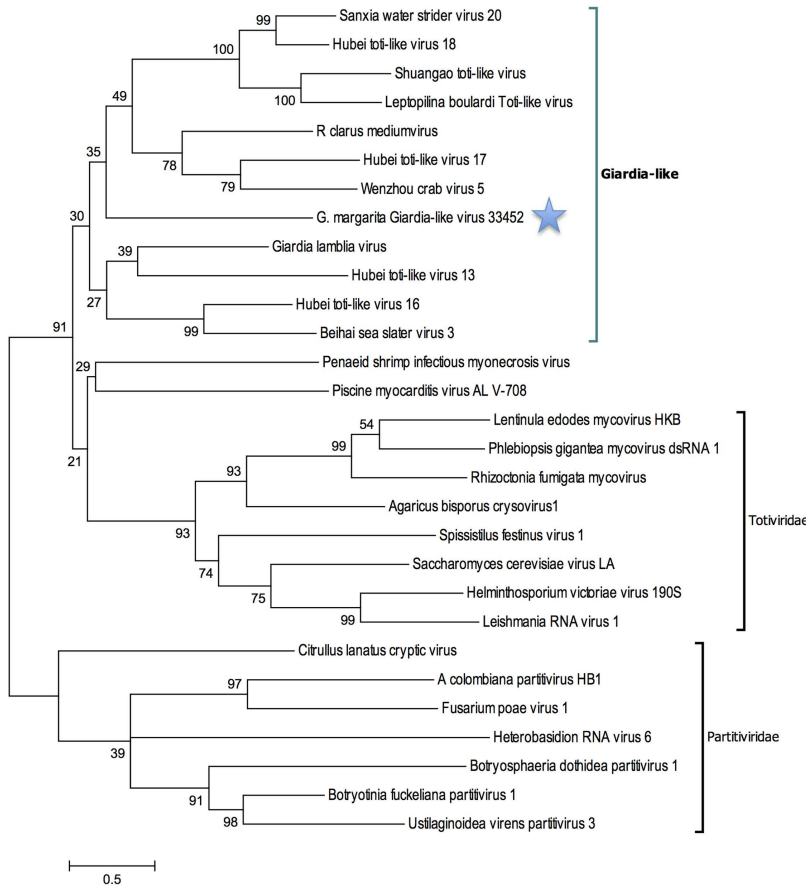


Fig. 4. Molecular Phylogenetic analysis by Maximum Likelihood method of the *Giardia*-like viral sequence in the context of some representative of the established family taxa *Totiviridae* and *Partitiviridae*. The tree with the highest log likelihood (-11851.6550) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 3.5294% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 29 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 170 positions in the final data set. The star indicates the position of the *G. margarita* *Giardia*-like virus characterized in this study.

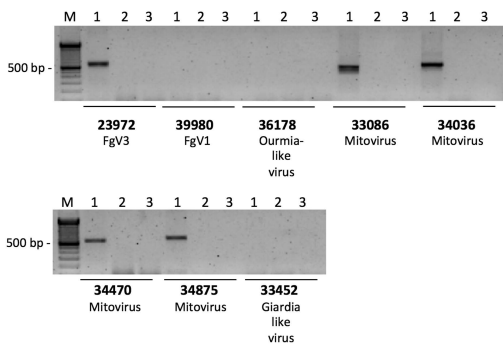


Fig. 5. Gel electrophoresis of PCR products of *G. margarita* DNA untreated (1) or treated (2) with DNase with primers specific for the viral sequences (the same primer pairs used in figure 1). Samples 3 correspond to reactions with no DNA. M: 100 bp DNA marker (Invitrogen).

assembly of *G. margarita* genome revealed the presence of the endogenized fragment in one scaffold (data not shown). Surprisingly, coverage was zero for the 4 mitoviral sequences, suggesting that a genomic integration was unlikely.

To investigate quantitatively the abundance of these viral DNA portions, we performed quantitative PCR assays using specific primers covering about 100–120 bp fragments of the four mitoviruses (Supporting Information Table S2) on three independent DNA preparations from 300 *G. margarita* spores. As endogenous controls we used PCR primers targeting a portion of the ribosomal DNA 18S and of the elongation factor gene. Only the DNA corresponding to the mitovirus 33086 accumulated to a level statistically significantly different from the other three mitoviruses. The amplicon corresponding to 33086 was the least abundant, while that of 34036 was the most abundant but in all cases the amount of viral DNA was statistically significantly lower than that of the DNA

Table 3. Ct values (\pm standard deviation) registered by qPCR on *G. margarita* DNA (3 independent samples) with primers for the 4 mitoviral sequences.

	33086	34086	34470	34875	Ef	18S rDNA
Ct	30.49	27.72	28.46	29.16	25.37	20.93
standard deviation	± 0.6647	± 0.2121	± 0.0141	± 0.1273	± 0.1697	± 0.3111
Tukey test	d	b	bc	c	a	e

18S rDNA and *Ef* (*Elongation factor*) gene were used as endogenous genes corresponding to multiple or single copy genes respectively. Different letters indicate statistically significant differences (ANOVA, $p < 0.05$).

sequence encoding the elongation factor, which is supposed to be a single copy gene (Table 3). Even with the more sensitive q-PCR assay we could not detect any DNA fragment related to the Ourmia-like sequence (data not shown). We cloned and sequenced the DNA fragments corresponding to the 34470 and 34875 mitoviruses, and confirmed that the sequence fragments had the same sequence as the one assembled *in silico*, without gaps or any signs of internal recombination (not shown).

To further clarify the nature of the DNA corresponding to the mitovirus sequences, we then investigated whether i) the full length DNA corresponding to the whole virus sequence could be amplified and ii) all the different regions of the viral genome were templates for DNA fragments. For this purpose a series of specific PCR primers was designed to span different portions of the genome of the contig 34036 and to obtain amplicons of different size for the same mitoviral sequence 34036 (Fig. 6): we selected for this experiment the mitovirus 34036 because it is the one with the highest DNA titer. The primers were also tested on two independent DNA extractions and on two independent cDNA preparations from germinating spores.

We could amplify relatively abundant DNA corresponding to fragments spanning the 3' terminal 2400 bp, and only a much fainter band corresponding to the fragment at the 5' of the genomic sequence (Fig 6A). The consensus sequence obtained from a number of different clones amplified from DNA corresponded exactly to that predicted from the *in silico* analysis, with only 1 synonymous nucleotide mutation (C to T) at position 1399 (Supporting Information Fig. S3). When we tested the maximum length of the amplified segment using different primer combination, independently from the genome fragments where primers were designed, only amplicons below 700 bp in length could be amplified from DNA, while larger PCR products were obtained from cDNA (Fig. 6B); when using the same DNA template in the same conditions, a 2 kb fragment spanning the elongation factor gene was obtained (not shown). This provides indirect evidence of the existence of small fragments of DNA corresponding to the various regions of most of the mitoviral genome.

The mitoviral sequences are present in *G. margarita* B- (*endobacteria free*) but not in *Gigaspora rosea* genomic DNA

To investigate the presence of homologous viral sequences in other AMF, we performed PCR experiments, using the *G. margarita* virome-specific primers (Supporting Information Table S2) on DNA from *Gigaspora rosea*, a phylogenetically related AMF, which does not host endobacteria. We also considered DNA of a *G. margarita* BEG34 isolate that has been cured from the *Candidatus* Glomeribacter gigasporarum endosymbiotic bacteria (Lumini *et al.*, 2007) and has been called B- (*endobacteria-free*). The quality of the DNA preparations was first verified by PCR amplifications with AM-specific ribosomal primers AML1 and AML2 (data not shown). A PCR product of the

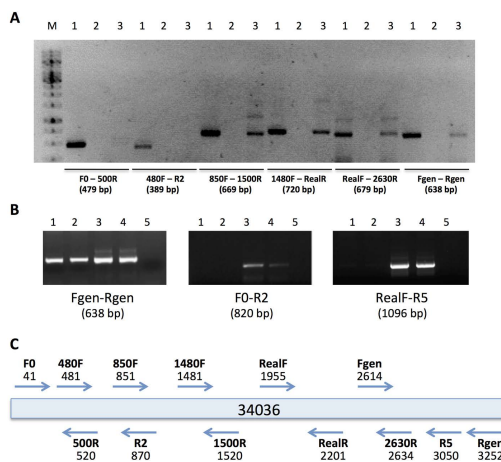


Fig. 6. Characters of the DNA corresponding to mitovirus sequences. A. Gel electrophoresis of PCR products obtained from *G. margarita* cDNA (1), DNA extractions after DNase (2) and DNA extractions (3). B. Gel electrophoresis of PCR products obtained from two *G. margarita* independent genomic DNA extractions (1, 2), two independent cDNAs preparations (3, 4) or no DNA (5) samples using the different primer pairs. The size of the amplicon in base pairs (bp) is indicated in brackets. C: Scheme of the 34 036 viral sequence showing the position of primers (F = forward; R = reverse) indicated by arrows.

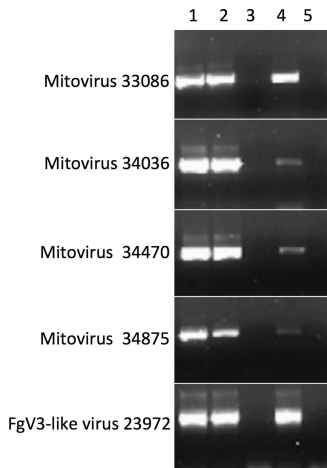


Fig. 7. Gel electrophoresis of PCR products from DNA samples: *G. margarita* B+ (1) and B- (2) harbouring endobacteria or endobacteria-cured respectively; *G. rosea* (3); *G. margarita* positive control (4); no DNA (5). The specific primer pair used for each virus is listed in the legend of Fig. 1.

expected size was obtained from the B- isolate for all the analysed viral sequences (4 mitoviruses and the FgV3-related sequence). The sequencing of the PCR products confirmed the correspondence to the expected virus sequence (not shown). By contrast, the *G. rosea* sample always gave negative results (Fig. 7). This suggests that these portions of viral DNA are constantly present in *G. margarita* independently from the presence of endobacteria, but are absent from the phylogenetically close *G. rosea*.

Abundance of viral RNA molecules in B+ and B- germinating spores

Since *Candidatus Glomeribacter gigasporarum* was shown to induce some physiological effects on the fungal host *G. margarita* (Salvioli *et al.*, 2016), we then investigated whether the abundance of the viral RNA molecules was regulated by the presence or absence of endobacteria. Quantitative RT-PCR reactions were set up on cDNA obtained from B+ and B- germinating spores using the elongation factor as housekeeping gene. On average, the mitoviral sequences led to a Ct (threshold cycle) between 11 and 14 (Ct = 12 for 33086; Ct = 14 for 34036; Ct = 11; 34470; Ct = 12 for 34875). For the Ouurmia-like virus and the Giardia-like virus sequences an average of 22 and 30 Ct was registered respectively, confirming the lower abundance compared to the mitoviral RNAs that was already noticed in the northern blot experiments. No statistically significant different virus accumulation between B+ and

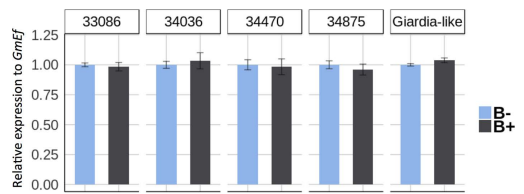


Fig. 8. Quantitative RT-PCR on cDNA of B+ and B- (harbouring endobacteria or endobacteria cured respectively) germinating spores of *G. margarita* with primers specific for the 4 mitovirus and the Giardia-like virus. *GmEf* (elongation factor) was used as housekeeping gene and B- sample as reference. No statistically significant difference was found between B+ and B- samples for any of the five viruses displayed (ANOVA; $p < 0.05$).

B- spores was observed for the mitoviral sequences and the Giardia-like virus (Fig. 8) but, surprisingly, the Ouurmia-like sequence seems to be present only in B+ spores: Ct values for this virus registered for B- spores were below the detection threshold.

Discussion

A further layer of complexity in the organisms associated to G. margarita

Recent works, such as the description of the genome sequence of *R. irregularis* (Tisserant *et al.*, 2013; Lin *et al.*, 2014), have increased our knowledge on AMF genetic and genomic complexity, and shed light on the molecular determinants that play a key role in the establishment of the symbiosis with the host plant. A further contribute to this complexity is given by the presence of endobacteria (Bonfante and Desirò, 2017), whose biological functions, with a single noticeable exception (Salvioli *et al.*, 2016), are still largely unknown.

A few studies recently revealed that a further level of complexity is given by the presence of mycoviruses, widespread symbionts in all major taxa of fungi. Mycoviruses in AMF have possibly evolved under unique selection pressures. In one case only it has been demonstrated that they are a biologically active component of the symbiosis: a *Glomus (Rhizophagus)* strain free of the GRF1V-M virus (a virus related to the Giardia-like virus described in this work) produced a higher amount of spores and promoted plant growth more efficiently than the fungal line containing the virus (Ikeda *et al.*, 2012). Similarly a fungal strain of *Mortierella elongata* cured from its endobacteria produces higher spore numbers and grows better (Li *et al.*, 2017). Nevertheless, virus and endobacterial association with AMF seem to be fairly common in nature, raising the possibility of some ecological advantage difficult to measure in the laboratory experimental conditions, where most of the intra- and interspecies competition aspects are not under scrutiny. Despite the above-mentioned major

acquisition, the knowledge of mycoviruses biological role in AMF is still limited.

In this work we identify, through computational analysis, four putative mitoviral sequences, a putative Ourmia-like mycovirus, a putative Giardia-like dsRNA virus and two sequences related to *F. graminearum* mycoviruses. The existence as RNA molecules of *in silico* predicted sequences was confirmed by RT-PCR experiments, for the 4 mitoviral sequences, and the Ourmia-like and Giardia-like sequences. We could not confirm the sequence related to FgV1 while that related to FgV3 turned out to be corresponding to a stably endogenized viral fragment into the genome of *G. margarita* likely not sufficiently transcribed to be detected by RT-PCR.

We could confirm the presence of the Ourmia-like virus RNA by RT-PCR assays only based on a short sequence of RNA molecule corresponding to a small portion of the sequence (Fig. 1B) which encodes for an RdRp conserved domain. We hypothesize that the whole 3.2 kb contig is probably a result of a misassembly. It is worthy to note that the contig is rich in AT stretches which often give many difficulties during the *in silico* assembly process. Nevertheless, the small segment corresponding to the part amplified by RT-PCR is indeed part of a self replicating RNA molecule, since no DNA corresponding to this fragment was detected, excluding endogenization events. Moreover, we demonstrated that the four mitoviruses are replicative and their RdRp is active since both positive and negative sense RNA strands were detected in northern blot assays. Regarding the Ourmia-like fragment, we obtained a signal only from the positive probe, while no signal was detected neither by the positive nor by the negative probe for the Giardia-like, probably due to low RNA abundance. The size of the Ourmia-like RNA detected with the antisense probe (detecting + sense) is approximately 3 kb: therefore indeed an RNA molecule larger than the small fragment we could amplify by RT-PCR exists.

Since many AMF are characterized by the presence of endobacteria which live inside hyphae and spores, and it has been demonstrated that their presence may have an important role for the fungal fitness (Lumini *et al.*, 2007; Salvioli *et al.*, 2016; Vannini *et al.*, 2016), we investigated whether the presence or absence of endobacteria could affect the occurrence and abundance of viral RNA in *G. margarita*. At least in germinating spores, the presence or absence of the endobacteria does not affect the occurrence and the abundance of viral RNA molecules, with the exception of the Ourmia-like virus that seems to be present only in B+ spores. Although our results are based on the analysis of a single cured-strain (the only available at this time), these findings open the possibility that the replication of this virus is dependent on endobacteria; indeed, a number of viruses phylogenetically related to Ourmia-like

viruses are known phages (family *Leviviridae*) (Dolja and Koonin, 2012).

Phylogenetic analysis of the G. margarita virome showed both evidence of virus-host co-evolution and of horizontal gene transfer

A phylogenetic analysis considering the conserved catalytic domain of the identified RdRps allowed us to associate the four mitoviruses to the genus *Mitovirus* and the Ourmia-like fragment to a new clade called Ourmia-like group that was recently shown to exist after a wide NGS (Next Generation Sequencing) analysis of invertebrate virome (Shi *et al.*, 2016).

A further analysis allowed us to focus on the diversity of the four mitoviruses; three of them cluster together (showing evidence of co-evolution with their host) while the 33086 sequence is separated and more closely related to mitoviruses from a phylogenetically distant host (*Botrytis cinerea*), providing some indirect evidence of possible horizontal virus transmission. Taxonomically, the criteria to establish a new mitovirus species requires the identity in the whole aligned RdRp to be below 40% (Hillman and Esteban, 2011): none of the four mitovirus found associated to *G. margarita* reached this threshold, implying that the four mitoviruses constitute new virus species respectively called *Gigaspora margarita* mitovirus 1 (GmMV1), GmMV2, GmMV3 and GmMV4.

The Giardia-like sequence also constitutes a new viral species called *G. margarita* Giardia-like virus 1 (GmGLV1). Given that in the case of the Ourmia-like sequence only a small fragment was confirmed by RT-PCR, we think that the data is overall too preliminary to establish a new virus species. The phylogenetic inference of the Giardia-like sequence shows that RdRp belonging to *Totiviridae* and *Partitiviridae* are grouped in a clade statistically supported while the other aligned sequences, including Giardia-like viruses, still do not form a statistically well supported clade (Fig. 4).

G. margarita mitoviruses display the same TGG codon frequency for tryptophan as the mitochondrial genes

The Mitovirus genus includes fungal viruses with small RNA genomes with a single ORF encoding a RdRp; most mitoviruses replicate in their host's mitochondria (Cole *et al.*, 2000; Hillman and Cai, 2013; Wu *et al.*, 2016). For translation, mitoviruses rely on the endogenous mitochondrial translational code which can use for tryptophan (Trp) the codons TGG (as in the cytosolic/nuclear genetic code) and TGA (which, by contrast, serves as a stop codon in the standard cytosolic/nuclear genetic code). Interestingly, in the four *G. margarita* mitoviral sequences all the Trp residues are encoded by the TGG codon. This observation

indicates that, apparently, functional RdRps from the four *G. margarita* mitoviruses can be translated both in cytosol and mitochondria. The use of TGG, compatible with both cytoplasmic and mitochondrial translation, is a feature of a few mitoviruses, including RcMV1-RF1, the only mitovirus described so far in an AMF (Kitahara *et al.*, 2014). It has been hypothesized that this feature might be an advantage for horizontal transmission among AMF (Kithara *et al.*, 2014). However, it has been recently observed that host fungi whose mitoviruses have no or few TGA codons are distinctive in also having no or few TGA codons in their core mitochondrial genes. For example, *G. margarita* mitochondrial DNA has only 2% Trp encoded by TGA and similar low percentages occur in the mitochondrial genomes of other AMF (Nibert, 2017). Thus, the exclusion of such codons in some mitoviruses appears to reflect most fundamentally that TGA is a rare mitochondrial codon in their particular hosts.

First report of cDNA corresponding to replicating non retroviral RNA viruses present in hosts outside the animal kingdom

Since integration events of fungal mitovirus cDNA in the mitochondrial (and to a minor extent nuclear) DNA of vascular plants was shown to be a fairly common event (Bruenn *et al.*, 2015), we investigated whether genome integration occurred also for the viral sequences identified in *G. margarita*. Our PCR assays revealed the presence of DNA fragments corresponding to the four mitoviruses and to FgV3-related sequence (23972), but when we checked the published sequence of the mitochondrial genome of *G. margarita* we did not find any viral sequence. The presence of multiple viral sequences in the same individual organism is not uncommon for mycoviruses (Kondo *et al.*, 2013; Xie and Ghabrial, 2012; Jiang *et al.*, 2013; Nerva *et al.*, 2016), even in the case of closely related viruses as for the 4 *G. margarita* mitovirus species. It would be interesting to assess the exact cellular niche for each mitovirus and whether they all infect each mitochondrion (mixed infection) or whether one mitovirus species for each mitochondrion exists, contributing to a differentiation among the mitochondrial lineages inside a single organism.

We had also the possibility to check for integration of those mitoviral DNA templates in the genomic DNA, looking for exact matches into reads from an ongoing sequencing project of the same *G. margarita* isolate: although the genome assembly is partial, we could map reads from the endogenized FgV3-like virus fragment, but no evidence of endogenized mitovirus sequence could be obtained. The nature of the amplified DNA for mitoviral sequences remains therefore undefined. Our attempts to perform Southern blot assays failed in detecting specific hybridization signals related to mitovirus endogenization

(data not shown). Taken as a whole, these data suggest that the mitoviral DNA sequences are not integrated in the genome but they likely exist as extrachromosomal fragments of relatively small size as it can be deduced by our amplicon-length PCR experiments. This is the first time that viral DNA has been detected in the presence of the corresponding homologous replicating RNA for mitoviruses. The presence of genome endogenization in plants has been so far associated with the absence of replicative capability of mitoviruses (Koonin and Dolja, 2014). To our knowledge this is also the first time that cDNA corresponding to a non retroviral replicating RNA virus is detected outside the animal kingdom. In fact only a few reports have pointed to the existence of episomic cDNA of RNA viruses in the order Diptera (Goic *et al.*, 2013; 2016; Nag *et al.*, 2016) and in the case of RNA viruses infecting mammals (Klenerman *et al.*, 1997; Weiss and Kellam, 1997). Recently, it was shown that two species of *Aedes* mosquitoes infected with two arboviruses from distinct families (dengue or chikungunya) generate, by endogenous reverse transcriptase activity, a viral-derived DNA that is essential for mosquito survival and viral tolerance, being at the base of persistent viral infections (Goic *et al.*, 2016); inhibition of cDNA synthesis results in higher mortality without affecting the RNAi anti-viral system. In our system, we have shown that cDNA sequences correspond exactly to viral genomic sequences without signs of re-arrangement, contrary to what shown for the *Drosophila melanogaster* viral infection system (Goic *et al.*, 2013). A possible analogy between the two system is that in mosquito cDNA was associated to persistent infection and the *G. margarita* virus-fungal system also has the hallmarks of persistent infections (lack of external infectivity of mitoviruses, widespread association between these viruses and AMF hosts, lack of obvious detrimental effects on the host, presence of a relatively high virus titer); it is tempting to speculate that the same anti-viral mechanism bringing to persistent infections is common to such distantly related hosts. Furthermore, it would be interesting to identify the mechanisms by which the mitoviral DNAs are generated in *G. margarita* and whether they may have an anti-viral functional role as is the case of mosquito infecting virus *via* the piRNA pathway (reviewed in Olson and Bonizzoni, 2017). Given the large amount of retrotransposon sequences present in the *G. margarita* genome (Ghignone, Venice, Anselem, Salvioli, Bonfante, unpublished), we can envision reverse transcriptase activity that can occasionally use as template RNA mitovirus sequences as demonstrated in other systems (Shimizu *et al.*, 2014); the presence of such retroviral activity in mitochondria is, to our knowledge, unknown and should be further tested.

In conclusions, the augmented genome concept that scientists often use to describe human beings at the organismal level, including the whole microbiota genomes,

can be in part applied to the fascinating biological system of AMF, where a fungus, its associated bacteria, and the associated viruses seem to constitute an interacting superorganism.

Experimental procedures

Biological material

Spores of *Gigaspora margarita* Becker and Hall (BEG 34), the corresponding cured strain (without the endobacterium *Candidatus Glomeribacter gigasporarum*; Lumini *et al.*, 2007) and *Gigaspora rosea* (BEG 9) were propagated using white clover (*Trifolium repens*) as trap plant. Clover plants were inoculated with ca. 100 spores and after 2–3 months new spores were generated and collected by the wet sieving technique. To generate germinating spores, *G. margarita* spore suspensions were divided into aliquots of 100, surface sterilized twice for 10 min with 3% chloramine-T and 0.03% streptomycin sulfate, rinsed several times with sterile distilled water and then incubated in 1 ml of sterile distilled water for 5–7 days in the dark at 30°C. Germinated spores were collected, immediately frozen in liquid nitrogen and stored at –80°C.

Total nucleic acids extraction

About 200–300 spores were crushed by a pestle using 1 ml of lysis CTAB buffer containing 2% cetyl trimethylammonium bromide, 1% polyvinyl pyrrolidone, 100 mM Tris-HCl, 1.4 M NaCl, 20 mM EDTA. After incubation at 65°C for 15 min, samples were centrifuged at 10 000 rpm for 10 min. The supernatant was added with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1; vol:vol:vol) and mixed. The upper aqueous phase of a centrifugation at 10 000 rpm for 10 min was transferred into a new tube to which an equal volume of chloroform was added. The aqueous phase was then collected and precipitated with 2/3 vol of isopropanol. After an incubation at 4°C for 2 h, samples were centrifuged at 14 000 for 30 min at 4°C. The pellet was washed with 80% ethanol and re-suspended into 50 µl of sterile water. To confirm that the PCR products were indeed originated from DNA, a control sample was generated by a DNase treatment performed on 1 µg of total nucleic acids for 1 h at 37°C using DNase (Ambion) according to manufacturer's instructions.

PCR and cloning

PCR reactions were set up following standard procedures. For sequencing the high fidelity Phusion DNA polymerase (ThermoFisher) was used. Primers are shown in Supporting Information Table S2. PCR products were cloned into the TOPO Vector (Invitrogen) vector following manufacturer's instructions. Recombinant DNAs were extracted with QIAGEN plasmid Minikit and sequences were obtained from the Sequencing Service, LMU Biozentrum, Großhaderner, Germany). Sequence analyses were performed with CHROMAS LITE (http://www.techneylum.com.au/chromas_lite.html).

RNA extractions and RT-PCR

Total RNA was extracted using the RNeasy Microarray Tissue Mini Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. The concentration and quality of the nucleic acids were assessed with a Nanodrop1000 (Thermo Scientific, Wilmington, NC, USA).

Samples were treated with TURBO DNase (Ambion) according to the manufacturer's instructions. The RNA samples were routinely checked for DNA contamination by means of RT-PCR analysis, using primers Efgig2F 5'-TGAACCT CCAACCAGACCAACTG-3' and EfgigR 5'-CGGTTTCAACA CGACCTACAGGGAC-3' for *G. margarita* elongation factor

(Efgig, Salvioli *et al.*, 2014) and the One-Step RT-PCR kit (Qiagen). For single-strand cDNA synthesis samples were denatured at 65°C for 5 min and then reverse-transcribed at 25°C for 10 min, 42°C for 50 min and 70° for 15 min in a final volume of 20 µl containing 10 µM random hexamers, 0.5 mM dNTPs, 4 µl 5X buffer, 2 µl 0.1 M DTT and 1 µl Super-ScriptIII (Invitrogen). Quantitative RT-PCR (qRT-PCR) experiments were carried out in a final volume of 10 µl containing 5 µl of iTaq™ Universal SYBR® Green Supermix (Bio-Rad), 0.2 µl of 2 µM specific primers (Table 2), and about 20 ng of cDNA. Samples were run in the iCycler iQ apparatus (Bio-Rad) using the following program: 3 min preincubation at 95°C, followed by 40 cycles of 10 s at 95°C, and 30 s at 60°C. Each amplification was followed by melting curve analysis (60–94°C) with a heating rate of 0.5°C every 15 s. All reactions were performed on at least three biological and three technical replicates and only Ct values with a standard deviation that did not exceed 0.3 were considered. To calculate primer efficiency standard dilution curves for each primer pair were set up; only primers pairs with similar efficiencies were used. The comparative threshold cycle method (Rasmussen, 2001) was used to calculate relative expression levels using as a reference gene for transcript normalization the *G. margarita* elongation factor (Efgig). Statistical analyses were performed through one-way analysis of variance (one-way ANOVA) and Tukey's *post hoc* test, using a probability level of $p < 0.05$. All statistical analyses were performed using the PAST statistical package (version 2.16; Hammer *et al.*, 2001).

Northern blot

For northern blot analyses, total RNA from *G. margarita* (BEG34) germinating spores and from MUT4330 (*Mycoteca Universitatis Taurinensis*, University of Torino), was prepared using Total Spectrum RNA Reagent (Sigma-Aldrich, Saint Louis, MO, USA) as suggested by the manufacturer. RNAs (about 2 µg/lane) were separated in gel electrophoresis under denaturing conditions using glyoxal in HEPES-EDTA buffer as detailed in Sambrook and colleagues (1987). Hybridization were performed using a radiolabelled RNA probe prepared from *EcoRI* linearized purified plasmids containing cDNA fragments corresponding to both orientation (sense and antisense probes) of each of the 6 viruses through T7 transcription using the Maxiscript T7 kit reagents (Thermo Fisher Scientific Inc., Waltham, MA, USA), as detailed before (Nerva *et al.*, 2017). In some cases, some fragments in specific orientations were toxic to *E. coli* after transformation, and recombinant plasmid could not be obtained. The problem was circumvented using as template for

transcription, instead of a linearized plasmid, a PCR fragment amplified with T7 promoter primer, and a fragment specific reverse primer directly amplified from the ligation reaction.

Phylogenetic analyses

We identified groups of conserved sequences related to the different mycoviruses present in *G. margarita* through BLAST searches of the databases (Accession numbers reported in Supporting Information Table S4) and the representative protein sequences identified were used for multiple sequence alignments using MUSCLE (Edgar, 2004) implemented in the MEGA 6 (Tamura *et al.*, 2013). Aligned sequences were used for infer phylogenetic trees using the Maximum Likelihood method based on the Le_Gascuel_2008 model (Le and Gascuel, 2008). The best amino acid substitution model was calculated with MEGA 6. Statistical analysis was carried out through bootstrap analysis with 1000 replicates. Further details of the phylogenetic analysis are included in figure legends. Multiple aligned sequences were also used to calculate pairwise identity and similarity percentages using MatGat (Campanella *et al.*, 2003).

Bioinformatic analyses

G. margarita BEG34 genomic DNA libraries (PE, MP-3kb, MP-8kb), currently used for an ongoing genome assembly project, were first checked for quality with FASTQC (Andrews, 2010) and then trimmed with TRIM GALORE! (Krueger, 2012). Cleaned PE reads were mapped onto *G. margarita* transcript comp11141_c0_seq1 (GBYF01010162.1), coding for Translation Elongation Factor EF-1 alpha, and onto putative viral sequences (33086, 34036, 34470, 34875) using BWA (Li and Durbin, 2009). Mapping outputs were handled and analysed with SAMTOOLS (Li *et al.*, 2009). Clues of endogenization of the putative FgV3 in the *G. margarita* genome where searched querying the assembly with the comp23972_c0_seq1 (GBYF01024012) sequence using BLASTN 2.6.0+ (Zhang *et al.*, 2000), and mapped reads were counted after BWA analysis.

The viral sequences have been submitted to GenBank under the following accession numbers: 33086: MG256173, 34036: MG256174, 34470: MG256175, 34875: MG256176 and 33452: MG256177.

Acknowledgements

The research was supported by the 60% Projects (University of Torino) to PB and LL. The Authors thank Dr. Mara Novero for the technical assistance and Dr. Simone Belmondo for its contribution in the first phases of the project.

References

Andrews, S. (2010) *FastQC: a quality control tool for high throughput sequence data* [WWW document]. URL: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
 Berruti, A., Lumini, E., Balestrini, R., and Bianciotto, V. (2015) Arbuscular mycorrhizal fungi as natural biofertilizers: let's benefit from past successes. *Front Microbiol* **6**: 1559.

Bonfante, P., and Desirò, A. (2017) Who lives in a fungus? The diversity, origins and functions of fungal endobacteria living in Mucoromycota. *Isme J* **11**: 1727–1735.
 Bruenn, J.A., Warner, B.E., and Yerramsetty, P. (2015) Widespread mitovirus sequences in plant genomes. *PeerJ* **3**: e876.
 Campanella, J.J., Bitincka, L., and Smalley, J. (2003) MatGAT: An application that generates similarity/identity matrices using protein or DNA sequences. *BMC Bioinform* **4**:29.
 Choi, G.H., and Nuss, D.L. (1992) A viral gene confers hypovirulence-associated traits to the chestnut blight fungus. *Embo J* **11**: 473–477.
 Cole, T.E., Hong, Y., Brasier, C.M., and Buck, K.W. (2000) Detection of an RNA-dependent RNA polymerase in mitochondria from a mitovirus-infected isolate of the Dutch Elm disease fungus, *Ophiostoma novo-ulmi*. *Virology* **268**: 239–243.
 Corradi, N., and Brachmann, A. (2017) Fungal mating in the most widespread plant symbiosis? *Trends in Plant Sci* **22**: 175–183.
 Desirò, A., Salvioli, A., Ngonkeu, E.L., Mondo, S.J., Epis, S., and Faccio, A. (2014) Detection of a novel intracellular microbiome hosted in arbuscular mycorrhizal fungi. *Isme J* **8**: 257–270.
 Dolja, V.V., and Koonin, E.V. (2012) Capsid-Less RNA Viruses. In: *eLS*. Chichester: John Wiley & Sons Ltd. URL: <http://www.els.net>
 Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.
 Ghabrial, S.A., and Suzuki, N. (2009) Viruses of plant pathogenic fungi. *Annu Rev Phytopathol* **47**: 353–384.
 Ghabrial, S.A., Castón, J.R., Jiang, D., Nibert, M.L., and Suzuki, N. (2015) 50-plus years of fungal viruses. *Virology* **479–480**: 356–368.
 Ghignone, S., Salvioli, A., Anca, I., Lumini, E., Ortu, G., and Petiti, L. (2012) The genome of the obligate endobacterium of an AM fungus reveals an interphyllum network of nutritional interactions. *Isme J* **6**: 136–145.
 Goic, B., Vodovar, N., Mondotte, J.A., Monot, C., Frangeul, L., Blanc, H., *et al.* (2013) RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model *Drosophila*. *Nat Immunol* **14**: 396–403.
 Goic, B., Stapleford, K.A., Frangeul, L., Doucet, A.J., Gausson, V., Blanc, H., *et al.* (2016) Virus-derived DNA drives mosquito vector tolerance to arboviral infection. *Nat Commun* **7**: 12410.
 Jiang, D., Fu, Y., Li, G., and Ghabrial, S.A. (2013) Viruses of the plant pathogenic fungus *Sclerotinia sclerotiorum*. *Adv Virus Res* **86**: 215–248.
 Hammer, Ø., Harper, D.A.T., and Ryan, P.D. (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4**(1): 9.
 Hillman, B.I., and Cai, G. (2013) The family Narnaviridae: simplest of RNA viruses. *Adv Virus Res* **86**: 149–176.
 Hillman, B.I., and Esteban, R. (2011) Narnaviridae. In: *Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses*. King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (eds). San Diego: Elsevier, pp.1055–1060.
 Hillman, B.I., Supyani, S., Kondo, H., and Suzuki, N. (2004) A reovirus of the fungus *Cryphonectria parasitica* that is

- infectious as particles and related to the coltivirus genus of animal pathogens. *J Virol* **78**: 892–898.
- Ikeda, Y., Shimura, H., Kitahara, R., Masuta, C., and Ezawa, T. (2012) A novel virus-like double-stranded RNA in an obligate biotroph arbuscular mycorrhizal fungus: a hidden player in mycorrhizal symbiosis. *Mol Plant-Microbe Interact* **25**: 1005–1012.
- Kamel, L., Keller-Pearson, M., Roux, C., and Ané, J.-M. (2017) Biology and evolution of arbuscular mycorrhizal symbiosis in the light of genomics. *New Phytol* **213**: 531–536.
- Kitahara, R., Ikeda, Y., Shimura, H., Masuta, C., and Ezawa, T. (2014) A unique mitovirus from Glomeromycota, the phylum of arbuscular mycorrhizal fungi. *Arch Virol* **159**: 2157–2160.
- Klenerman, P., Hengartner, H., and Zinkernagel, R.M. (1997) A non-retroviral RNA virus persists in DNA form. *Nature* **390**: 298–301.
- Kondo, H., Kanematsu, S., and Suzuki, N. (2013) Viruses of the white root rot fungus, *Rosellinia necatrix*. *Adv Virus Res* **86**: 177–214.
- Koonin, E.V., and Dolja, V.V. (2014) Virus world as an evolutionary network of viruses and capsidless selfish elements. *Microbiol Mol Biol Rev* **78**: 278–303.
- Krueger, F. (2012) A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced Representation Bisulfite-Seq) libraries. URL: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
- Le, S.Q., and Gascuel, O. (2008) An improved general amino acid replacement matrix. *Mol Biol Evol* **25**: 1307–1320.
- Liu, S., Xie, J., Cheng, J., Li, B., Chen, T., Fu, Y., et al. (2016) Fungal DNA virus infects a mycophagous insect and utilizes it as a transmission vector. *Proc Natl Acad Sci USA* **113**: 12803–12808.
- Li, H., and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* **25**: 1754–1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009) The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* **25**: 2078–2079.
- Li, Z., Yao, Q.M., Dearth, S.P., Entler, M.R., Castro Gonzalez, H.F., Uehling, J.K., et al. (2017) Integrated proteomics and metabolomics suggests symbiotic metabolism and multimodal regulation in a fungal-endobacterial system. *Env Microb* **19**: 1041–1053.
- Lin, K., Limpens, E., Zhang, Z.H., Ivanov, S., Saunders, D.G.O., Mu, D.S., et al. (2014) Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. *PLoS Genet* **10**: e1004078.
- Lumini, E., Bianciotto, V., Jargeat, P., Novero, M., Salvioli, A., Faccio, A., et al. (2007) Presymbiotic growth and spore morphology are affected in the arbuscular mycorrhizal fungus *Gigaspora margarita* cured of its endobacteria. *Cellular Microbiol* **9**: 1716–1729.
- Márquez, L.M., Redman, R.S., Rodriguez, R.J., and Roossinck, M.J. (2007) A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. *Science* **315**: 513–515.
- Nag, D.K., Brecher, M., and Kramer, L.D. (2016) DNA forms of arboviral RNA genomes are generated following infection in mosquito cell cultures. *Virology* **498**: 164–171.
- Naumann, M., Schübler, A., and Bonfante, P. (2010) The obligate endobacteria of arbuscular mycorrhizal fungi are ancient heritable components related to the Mollicutes. *Isme J* **4**: 862–871.
- Nerva, L., Ciuffo, M., Vallino, M., Margaria, P., Varese, G.C., Gnani, G., and Turina, M. (2016) Multiple approaches for the detection and characterization of viral and plasmid symbionts from a collection of marine fungi. *Virus Res* **219**: 22–38.
- Nerva, L., Silvestri, A., Ciuffo, M., Palmano, S., Varese, G.C., and Turina, M. (2017) Transmission of *Penicillium aurantiogriseum* partiti-like virus 1 to a new fungal host (*Cryphonectria parasitica*) confers higher resistance to salinity and reveals adaptive genomic changes. *Environ Microb* **19**: 4480–4492.
- Nibert, M.L. (2017) Mitovirus UGA(Trp) codon usage parallels that of host mitochondria. *Virology* **507**: 96–100.
- Olson, K.E., and Bonizzoni, M. (2017) Nonretroviral integrated RNA viruses in arthropod vectors: an occasional event or something more?. *Curr Opin Insect Sci* **22**: 45–53.
- Pelin, A., Pombert, J.F., Salvioli, A., Bonen, L., Bonfante, P., and Corradi, N. (2012) The mitochondrial genome of the arbuscular mycorrhizal fungus *Gigaspora margarita* reveals two unsuspected transsplicing events of group I introns. *New Phytol* **194**: 836–845.
- Rasmussen, R. (2001) Quantification on the LightCycler. In *Rapid Cycle Real-Time PCR: Methods and Applications*. Mener, S., Wittwer, C., Nakagawara, K. (eds). Heidelberg: Springer Press, pp. 21–34.
- Rodriguez, A., and Sanders, I.R. (2015) The role of community and population ecology in applying mycorrhizal fungi for improved food security. *Isme J* **9**: 1053–1061.
- Ropars, J., Toro, K.S., Noel, J., Pelin, A., Charron, P., Farinelli, L., et al. (2016) Evidence for the sexual origin of heterokaryosis in arbuscular mycorrhizal fungi. *Nat Microbiol* **1**: 16033.
- Salvioli, A., Ghignone, S., Novero, M., Navazio, L., Venice, F., Bagnaresi, P., and Bonfante, P. (2016) Symbiosis with an endobacterium increases the fitness of a mycorrhizal fungus, raising its bioenergetic potential. *Isme J* **10**: 130–144.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1987) *Molecular Cloning: A Laboratory Manual*. 2nd ed. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press.
- Sędziewlewska Toro, K., and Brachmann, A. (2016) The effector candidate repertoire of the arbuscular mycorrhizal fungus *Rhizophagus clarus*. *BMC Genomics* **17**: 101.
- Shi, M., Lin, X.D., Tian, J.H., Chen, L.J., Chen, X., Li, C.X., et al. (2016) Redefining the invertebrate RNA virosphere. *Nature* **540**: 539–543.
- Shimizu, A., Nakatani, Y., Nakamura, T., Jinno-Oue, A., Ishikawa, O., Boeke, J.D., et al. (2014) Characterisation of cytoplasmic DNA complementary to non-retroviral RNA viruses in human cells. *Sci Rep* **4**: 5074.
- Son, M., Yu, J., and Kim, K.-H. (2015) Five questions about mycoviruses. *PLoS Pathog* **11**: e1005172.
- Spatafora, J.W., Chang, Y., Benny, G.L., Lazarus, K., Smith, M.E., Berbee, M.L., et al. (2016) A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* **108**: 1028–1046.
- Stielow, B., and Menzel, W. (2010) Complete nucleotide sequence of TaV1, a novel totivirus isolated from a black truffle ascocarp (*Tuber aestivum* Vittad.). *Arch Virol* **155**: 2075–2078.

- Stielow, B., Klenk, H.P., Winter, S., and Menzel, W. (2011a) A novel *Tuber aestivum* (Vittad.) mitovirus. *Arch Virol* **156**: 1107–1110.
- Stielow, B., Klenk, H.P., Winter, S., and Menzel, W. (2011b) Complete genome sequence of the first endornavirus from the ascocarp of the ectomycorrhizal fungus *Tuber aestivum* Vittad. *Arch Virol* **156**: 343–345.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., and Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**: 2725–2729.
- Tang, N., San Clemente, H., Roy, S., Bécard, G., Zhao, B., and Roux, C. (2016) A survey of the gene repertoire of *Gigaspora rosea* unravels conserved features among Glomeromycota for obligate biotrophy. *Front Microbiol* **7**: 233.
- Tisserant, E., Malbreil, M., Kuo, A., Kohler, A., Symeonidi, A., Balestrini, R., et al. (2013) Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proc Natl Acad Sci USA* **110**: 20117–20122.
- Torres-Cortés, G., Ghignone, S., Bonfante, P., and Schüßler, A. (2015) Mosaic genome of endobacteria in arbuscular mycorrhizal fungi: transkingdom gene transfer in an ancient mycoplasma-fungus association. *Proc Natl Acad Sci USA* **112**: 7785–7790.
- Turina, M., Hillman, B.I., Izadpanah, K., Rastgou, M., Rosa, C., and Consortium, I.R. (2017) ICTV Virus Taxonomy Profile: Ourmiavirus. *J Gen Virol* **98**: 129–130.
- van Diepeningen, A.D., Debets, A.J., and Hoekstra, R.F. (2006) Dynamics of dsRNA mycoviruses in black Aspergillus populations. *Fungal Genet Biol* **43**: 446–452.
- Vannini, C., Carpentieri, A., Salvioli, A., Novero, M., Marsoni, M., Testa, L., et al. (2016) An interdomain network: the endobacterium of a mycorrhizal fungus promotes antioxidative responses in both fungal and plant hosts. *New Phytol* **211**: 265–275.
- Weiss, R.A., and Kellam, P. (1997) Illicit viral DNA. *Nature* **390**: 235–236.
- Wu, M., Deng, Y., Zhou, Z., He, G., Chen, W., and Li, G. (2016) Characterization of three mycoviruses coinfecting the plant pathogenic fungus *Sclerotinia nivalis*. *Virus Res* **223**: 28–38.
- Xie, J., and Ghabrial, S.A. (2012) Molecular characterizations of two mitoviruses co-infecting a hypovirulent isolate of the plant pathogenic fungus *Sclerotinia sclerotiorum*. *Virology* **428**: 77–85.
- Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000) A greedy algorithm for aligning DNA sequences. *J Comput Biol* **7**: 203–214.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Nucleotide sequences of the full-length contigs assembled *in silico* corresponding to the virus sequences in fasta format. RdRp start and stop codons are indicated in

bold and cyan. When relevant, the corresponding translation frame is given to outline the codons encoding for tryptophan (W highlighted in yellow). Amino acid sequence of the RdRp is also then reported in fasta format and the GDD conserved motif is indicated in pink.

Fig. S2. Molecular Phylogenetic analysis by Maximum Likelihood method of the four mitoviruses and the Ourmia virus (positions in the tree pointed by red arrows). The evolutionary history was inferred using the Maximum Likelihood method based on the Le_Gascuel_2008 model (Le and Gascuel, 2008). The tree with the highest log likelihood (–11687.6554) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories [+G, parameter = 1.0325]). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 2.9115% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 64 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 109 positions in the final data set. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

Fig. S3. Schematic representation of the assembly of the clones used for deriving the 2kb consensus sequence corresponding to 34036 mitovirus sequence amplified by PCR using as template DNA preparations (compare with Fig. 6). In blue the amplified genomic region. Vertical green bars correspond to single nucleotide mutations in specific position in at least one of the cloned amplicons in respect to the consensus sequence. The only conserved mutation compared to the *in silico* assembled sequence from transcriptome is a synonymous C to T change at position 1399. Gray arrows represent the full-length viral contig sequence (derived from transcriptome, top arrow) and all the different sequences derived from cloning the PCR amplification products displayed in Fig. 6.

Table S1. BLAST searches of databases using as query the viral RdRp contigs present in *G. margarita* transcriptome.

Table S2. List of primers used in this study.

Table S3. Percentage identity (upper right triangle) and similarity (lower left triangle) among aligned RdRp from different mitoviruses.

Table S4. Accession numbers and virus names of proteins used in the phylogenetic analysis.

CHAPTER 5

5. Different genetic sources contribute to the small RNA population in the arbuscular mycorrhizal fungus *Gigaspora margarita*

Alessandro Silvestri¹, Valentina Fiorilli¹, Massimo Turina², Laura Miozzi²,
Francesco Venice¹, Paola Bonfante¹, Luisa Lanfranco¹

¹Department of Life Sciences and Systems Biology, University of Torino, Torino, Italy

²Institute for Sustainable Plant Protection, CNR, Torino, Italy.

Accepted on *Frontiers in Microbiology*

5.1. Abstract

RNA interference (RNAi) is a key regulatory pathway of gene expression in almost all eukaryotes. This mechanism relies on short non-coding RNA molecules (sRNAs) to recognize in a sequence-specific manner DNA or RNA targets leading to transcriptional or post-transcriptional gene silencing. To date, the fundamental role of sRNAs in the regulation of development, stress responses, defense against viruses and mobile elements and cross-kingdom interactions has been extensively studied in a number of biological systems. However, the knowledge of the “RNAi world” in arbuscular mycorrhizal fungi (AMF) is still limited. AMF are obligate mutualistic endosymbionts of plants, able to provide several benefits to their partners, from improved mineral nutrition to stress tolerance. Here we described the RNAi-related genes of the AMF *Gigaspora margarita* and characterized, through sRNA sequencing, its complex small RNAome, considering the possible genetic sources and targets of the sRNAs. *G. margarita* indeed is a mosaic of different genomes since it hosts endobacteria, RNA viruses and non-integrated DNA fragments corresponding to mitovirus sequences.

Our findings show that *G. margarita* is equipped with a complete set of RNAi-related genes characterized by the expansion of the *Argonaute-like* (*AGO-like*) gene family that seems a common trait of AMF. With regards to sRNAs, we detected populations of sRNA reads mapping to nuclear, mitochondrial and viral genomes that share similar features (25-nt long and 5'-end uracil read enrichments), and that clearly differ from sRNAs of endobacterial origin. Furthermore, the annotation of nuclear loci producing sRNAs suggests the occurrence of different sRNA-generating processes. *In silico* analyses indicate that the most abundant *G. margarita* sRNAs, including those of viral origin, could target transcripts in the host plant, through a hypothetical cross-kingdom RNAi.

Keywords: *Gigaspora margarita*, arbuscular mycorrhizal fungi, small RNA, RNA interference, viruses, symbiosis

5.2. Introduction

RNA interference (RNAi) or RNA silencing is a conserved eukaryotic pathway involved in the repression of gene expression at transcriptional or post-transcriptional level (Moazed, 2009; Wilson and Doudna, 2013; Ipsaro

and Joshua-Tor, 2015). RNAi carries out several biological functions, such as gene regulation, defense against mobile repetitive DNA sequences, retroelements, transposons and viruses. The process is mediated by small RNAs (sRNAs) of about 20-30 nucleotides that direct, by sequence complementarity, the recognition and silencing of the target genetic elements. The RNAi pathway relies on three core enzymes: Dicer-like (DCL), Argonaute-like (AGO-like) and RNA-dependent RNA polymerase (RdRp). DCL are ribonuclease III (RNase III) proteins that cleave double-stranded RNAs (dsRNAs) or single-stranded hairpin RNAs producing sRNAs that are then loaded onto AGO-like which, guided by sRNAs, are responsible for the silencing of the specific target sequences. RdRp generally play a dual role, both triggering the RNAi pathway and/or amplifying the silencing signals through the synthesis of dsRNAs from aberrant RNAs.

In 2013, Weiberg et al. (2013) discovered that RNAi is also a key molecular component of interspecies communication: sRNAs can be transferred across the contact interface of two interacting organisms and, acting as pathogen effectors, they silence specific genes in host cells in order to favor colonization. This phenomenon, known as cross-kingdom RNAi, occurs in several pathogenic and parasitic interactions (Weiberg et al., 2013; Mayoral et al., 2014; Zhang et al., 2016; Shahid et al., 2018; Chow et al., 2019; Cui et al., 2019) where it can function as an attack or a defense strategy. Interestingly, it was also recently described in the legume-rhizobium symbiosis where bacterial transfer RNA (tRNA)-derived small RNA fragments are signal molecules that modulate host gene expression and nodule formation (Ren et al., 2019). It has been proposed that sRNAs can also be exchanged between the partners of the arbuscular mycorrhizal (AM) symbiosis (Huang et al., 2019), a very ancient mutualistic association established between the roots of most plants and the obligate biotrophic fungi belonging to Glomeromycotina (Mucoromycota phylum; Spatafora et al., 2016), known as arbuscular mycorrhizal fungi (AMF) (Lanfranco et al., 2018). Indeed, host-induced (HIGS) and virus-induced gene silencing (VIGS) have been successfully employed to silence AMF genes expressed during root colonization (Helber et al., 2011; Kikuchi et al., 2016; Tsuzuki et al., 2016; Xie et al., 2016; Voß et al., 2018).

The RNAi machinery and the sRNA populations in AMF have been characterized only in the model AMF species *Rhizophagus irregularis* (Lee et al. 2018; Silvestri et al., 2019). The availability of the full genome sequence of

Gigaspora margarita (Venice et al., 2019) offered the possibility to explore how conserved are RNAi features in AMF. *G. margarita* is of particular interest as it belongs to Gigasporaceae, an early diverging AMF group, well separated from Glomeraceae that includes *R. irregularis* (Krüger et al., 2012). It has a complex genomic structure with the largest fungal genome so far annotated (773 Mbp) and a rich content (64%) in transposable elements. In addition, the isolate used for the genome project (BEG34) can be treated as a meta-organism, since it hosts the obligate endobacterium *Candidatus* Glomeribacter gigasporarum (CaGg; Ghignone et al., 2012) and six viral species (Turina et al., 2018), whose genome sequences are available. Notably, for the four mitoviruses present in *G. margarita* we could also prove the existence of DNA fragments corresponding to portions of their genome: this feature, never found in mycoviruses, has been described as an anti-viral response in insects (Goic et al., 2016).

In this work we described the RNAi-related gene components in *G. margarita* and the sRNA population originating from its metagenome. *G. margarita*, in analogy to *R. irregularis*, is equipped with a complete set of RNAi-related genes, characterized by the expansion of the AGO-like family, as well as sRNAs. A population of nuclear DNA mapping sRNAs substantially different from that of *R. irregularis* was found. The high level of sRNA reads mapping to viral genomes suggests that *G. margarita* RNAi machinery is able to provide an antiviral defense. Furthermore, through an *in silico* analysis, we identified a group of plant genes that can be potentially targeted by the most expressed *G. margarita* sRNAs.

5.3. Methods

Biological material

The spores of *G. margarita* strain BEG34 were obtained from *Trifolium repens* plants inoculated with 100–150 spores. After 3 months of growth with night/day temperature conditions of 21°C (night) and 23°C (day), new spores were collected by wet sieving technique, divided in batches of 100 and vernalized in distilled water for a week in the dark at 4°C. Spores were then surface-sterilized with chloramine T (3% W/V) and streptomycin sulfate (0.03% W/V), washed with sterile distilled water and incubated in 1 ml of sterile distilled water for a week in the dark at 30°C to allow germination. Finally,

germinated spores were collected, immediately frozen in liquid nitrogen, lyophilized and stored at -80°C .

RNA extraction for sRNA-seq

Batches of 100 germinated spores were ground in a bead beater with 3-mm tungsten beads at 18 Hz/s for 3 min. Total RNA was extracted with Direct-zol™ RNA MiniPrep (Zymo Research) kit, performing the in column DNase I treatment as recommended by the manufacturer. RNA concentration and quality were assessed with a Nanodrop1000 (Thermo Scientific). Three biological replicates (SG1, SG2, SG3) were prepared pooling together, for each replicate, same amount of RNA extracted from 2 independent biological samples (so a total of 6 biological samples were used). Samples were then delivered to Macrogen (South Korea) for RNA integrity check, library preparations and sequencing.

Identification and phylogenetic analyses of RdRp, DCL and AGO-like

Screening the Pfam annotation of the *G. margarita* proteome (Venice et al., 2019), we retrieved all the sequences containing a “Piwi” (Pfam ID: PF02171.17), an “RdRP” (Pfam ID: PF05183.12) or two “Ribonuclease_3” (Pfam ID: PF00636.26) domains, and we considered them as AGO-like, RdRp or DCL, respectively. The whole amino acid sequences of DCL, AGO-like and RdRp were aligned with MAFFT v7.310 (option: --auto) (Kato and Standley, 2013) together with fungal sequences analyzed in Silvestri et al. (2019). Their phylogenetic relationships were inferred by the Maximum Likelihood method implemented in the IQ-TREE software (options: -m TEST -bb 1000 -alrt 1000 -o “root”) (Nguyen et al., 2015). The software performed model selection (Kalyaanamoorthy et al., 2017), tree reconstruction and branch support analysis by ultra-fast bootstrap method (1000 replicates) (Hoang et al., 2018). Trees were reshaped on root with Newick Utilities v1.6 (command: nw_reroot) (Junier and Zdobnov, 2010) and visualized with Evolview v3 (Subramanian et al., 2019).

Bioinformatics pipeline

Raw sRNA-seq reads were checked for quality with FastQC (Babraham Bioinformatics) and then cleaned from adapters (TGGAATTCTCGGGTGCCAAGG), artifacts (default parameters) and low quality

reads (-q 28 -p 50) with Fastx Toolkit (Hannon Lab). Further filtering of raw reads was performed with Bowtie (Langmead et al., 2009), by removing the reads that mapped with up to 1 mismatch to the tRNA, rRNA, snRNA and snoRNA sequences from Rfam 12.0 database (Nawrocki et al., 2015), and those mapping with 0 mismatches to all the “ribosomal RNA” sequences present in GenBank for Mucoromycota taxonomy; we finally kept only the 18- to 35-nt long reads. Filtered reads were mapped with no mismatch with bowtie to the unmasked versions of *G. margarita* nuclear (Venice et al., 2019), mitochondrial, (Pelín et al., 2012), CaGg (Ghignone et al., 2012) and viral (Turina et al., 2018) genomes. A set of Bash, Perl and R scripts were used for the analysis and visualization of nucleotide length distribution, 5'-end nucleotide composition and reads redundancy. For the analysis of nucleotide length distribution of sRNAs in *R. irregularis*, reads from “extra-radical mycelium” libraries of a previous study (Silvestri et al., 2019) were mapped to mitochondrial (NCBI accession: JQ514224.2) and nuclear genomes (Chen et al., 2018) of *R. irregularis* with no mismatch using bowtie.

ShortStack v.3.8.5 (Johnson et al., 2016) was used for genome-guided sRNA-generating loci prediction and annotation on *G. margarita* nuclear genome (options: --mismatches 0 --foldsize 1000 --dicermin 18 --dicermax 35 --pad 200 --mincov 10.0rpm). BEDTools (Quinlan and Hall, 2010) was used to compare the genomic locations of sRNA-generating loci with those of annotated protein-encoding genes and of annotated transposable elements (Venice et al., 2019). PCA on sRNA-generating loci was performed in R with “FactoMineR” v1.42 (Lê et al., 2008) and “factoextra” v1.0.5 (Kassambara and Mundt 2017) packages. HDBSCAN clustering was performed (parameters: minPts = 20) with dbscan R package v1.1-4 (Hahsler M, Piekenbrock M, 2019). Homology analysis of *G. margarita*-(*Gma*)-sRNA-generating loci with fungal repetitive elements from RepBase 23.04 (<https://www.girinst.org/replib/>) was performed with tblastx (E-value <= 0.00005) (Camacho et al., 2009). The miRNA-like locus was annotated by ShortStack v. 3.8.5 (Johnson et al., 2016) and its secondary structure was predicted and visualized with StrucVis v.0.3 (<https://github.com/MikeAxtell/strucVis>).

For target prediction analysis, we selected the 21-nt long sRNAs with expression level greater than 100 RPM (“Reads Per Million mapped reads”; considering only the 21-24-nt long reads mapped to *G. margarita* genome)

and the 21-nt viral sRNAs with expression level greater than 100 RPM among all viral sRNAs. These sRNA reads were then used to predict targets in *M. truncatula* A17 transcriptome (v.4.0 cDNAs on EnsemblPlants database; Zerbino et al., 2018) through psRNAtarget (2017 update; Dai et al., 2018) with default parameters and we kept only the predictions with expectation lower than 3. GO enrichment analysis of target transcripts was performed with AgriGO (p-value < 0.01; statistical test: Fisher's test with Yekutieli correction; Tian et al., 2017), using Plant Go Slim ontology. Similarly, two further target prediction analyses on *M. truncatula* transcriptome were performed as described above using only the most abundant 21-nt long sRNAs of the AMF *R. irregularis* (> 100 RPM, considering only the 21-24-nt long reads mapping on the genome) from "mycorrhizal roots" libraries described in Silvestri et al. (2019) and the most abundant 21-nt long sRNAs of the non-AMF *Aspergillus fumigatus* (> 100 RPM before genome mapping; Özkan et al., 2017). To identify the *A. fumigatus* sRNA sequences, two already published sRNA-seq libraries (SRA ID: SRR1583955, SRR1583956; Özkan et al., 2017) were cleaned from adapters (AGATCGGAAGAGCACACGTCT), artifacts, low qualities reads and tRNA-, rRNA-, snRNA- and snoRNA-related sequences as described above. The remaining reads were mapped with no mismatch with bowtie to the unmasked version of *A. fumigatus* Af293 genome from Ensembl database (ASM265v1). We kept, for the target prediction on the *M. truncatula* transcriptome, only the 21-nt long genome mapped sRNAs with an abundance greater than 100 "reads per million reads" of the whole filtered sRNA libraries.

5.4. Results and Discussion

The comparative analysis of fungal RNAi-related proteins reveals common AMF traits

Recent surveys of two AMF genomes allowed a first characterization of the RNAi components (*DCL*, *AGO-like* and *RdRp*) in this group of obligate biotrophs (Lee et al., 2018; Silvestri et al., 2019). These works highlighted that the two analyzed AMF, *R. irregularis* and *R. clarus*, are equipped with a RNAi machinery characterized by 26-40 AGO-like, 3-21 RdRp and 1-2 DCL proteins. Here we analyzed the genome of the AMF *G. margarita* (Venice et al.,

2019) in order to define the conservation level of the RNAi-related genes in the Glomeromycotina subphylum. Keeping only one virtual transcript for each gene, we obtained a total of 11 AGO-like, 6 RdRp and 1 DCL corresponding proteins.

In analogy to the other analyzed AMF, *R. irregularis* and *R. clarus*, *G. margarita* is characterized by an expansion, even if less pronounced compared to both *Rhizophagus* species, of the AGO-like gene family. Other filamentous fungi in the Ascomycota and Basidiomycota, in fact, typically possess 1-4 AGO (Chang et al., 2012). Only 5 out of the 11 *G. margarita* AGO-like proteins (g74.t1, g25280.t1, g11769.t1, g13419.t1, g17397.t1) show all the typical AGO domains (piwi, PAZ, MID and N-terminal) (Poulsen et al., 2013); the remaining 6 (g16172.t1, g19042.t1, g19043.t1, g24476.t1, g19778.t1, g20012.t1) lack at least one of the non-piwi domains (Figure 1). We can speculate that those atypical AGO-like can play different biological functions, unrelated to the classic RNAi pathway. A phylogenetic analysis of fungal AGO-like revealed that 10 out of 11 *G. margarita* proteins (g74.t1, g16172.t1, g19042.t1, g19043.t1, g24476.t1, g19778.t1, g20012.t1, g25280.t1, g11769.t1, g13419.t1) belong to a well-supported clade containing only proteins from Mucoromycota species (*G. margarita*, *Mucor circinnelloides* and *R. irregularis*; Mucoromycota-specific clade) while the remaining one (g17397.t1) groups with some AGO from Ascomycota (*Cryphonectria parasitica*, *Neurospora crassa* and *Magnaporthe oryzae*) and 5 *R. irregularis* AGO-like (non-Mucoromycota specific clade; Figure 1). Interestingly, the Mucoromycota-specific clade can be further divided in two subgroups, one that is AMF specific (*G. margarita* and *R. irregularis*). A third group, non-related to the previous two, contains only some Ascomycota AGO-like proteins (*C. parasitica*, *M. oryzae*, *N. crassa*, *Schizosaccharomyces pombe*; Ascomycota-specific clade). Moreover, *G. margarita* is not equipped with small peptide-encoding ORF containing only the piwi domain that were found in *R. irregularis* but not in *R. clarus* (Silvestri et al., 2019). Despite the different number of AGO-like, *R. irregularis* and *G. margarita* are equipped with the same core set of homologous sequences: *G. margarita* possesses at least one AGO-like protein for each phylogenetic subgroup present in *R. irregularis*.

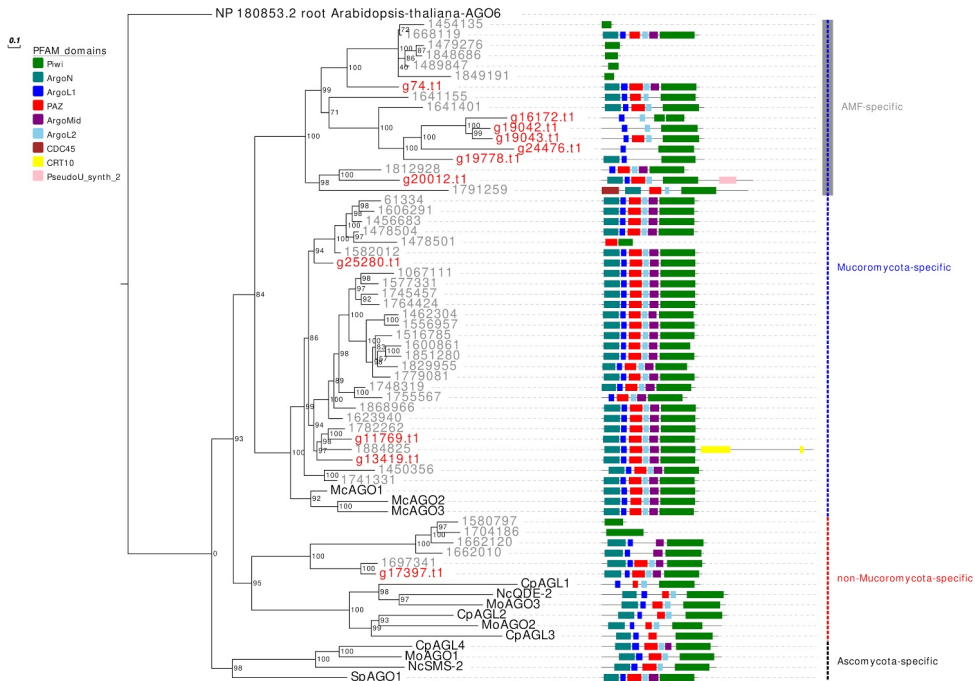


Figure 1. Phylogenetic analysis and PFAM domains of AGO-like proteins. Sequences are discernible by species according to a two-letter prefix and/or a color code: Mo = *Magnaporthe oryzae*, Nc = *Neurospora crassa*, Mc = *Mucor circinelloides*, Sp = *Schizosaccharomyces pombe*, Cp = *Cryphonectria parasitica*, grey = *Rhizophagus irregularis*, red = *Gigaspora margarita*. Protein ID (NCBI or JGI): MoAGO1 = XP_003716704.1, MoAGO2 = XP_003717504.1, MoAGO3 = XP_003714217.1, NcQDE-2 = XP_011394903.1, NcSMS-2 = EAA29350.1, SpAGO1 = O74957.1, McAGO-1 = 104,161, McAGO-2 = 195,366, McAGO-3 = 104,163, CpAGL1 = ACY36939.1, CpAGL2 = ACY36940.1, CpAGL3 = ACY36941.1, CpAGL4 = ACY36942.1. *R. irregularis* proteins are identified by JGI numeric codes. *G. margarita* proteins are identified by Venice et al. (2019) annotation code. The numbers at the nodes are bootstrap values (%) for 1000 replications. Tree was rooted using *Arabidopsis thaliana* Argonaute 6 (NCBI Reference Sequence: NP_180853.2). Tree was reshaped on root with Newick Utilities v1.6 and visualized with Evolvview v3.

With regards to the RdRp, we found 6 proteins in *G. margarita*, more than the 1-5 generally possessed by non-AMF (Chang et al., 2012; Chen et al., 2015). In our previous work we reported 21 RdRp in *R. irregularis*, although Lee et al. (2018), using more stringent annotation criteria, only reported 3.

The RdRp phylogenetic analysis revealed the presence of three main clades, each containing at least one protein sequence from *G. margarita* (Figure 2). A

first clade includes, considering only the Mucoromycota, 2 *G. margarita* (g20441.t1, g4496.t1), 15 *R. irregularis* and 2 *M. circinnelloides* sequences, which are related to the *N. crassa* SAD-1. The second clade contains a single *G. margarita* (g5332.t1) and 3 *R. irregularis* sequences, which are related to *N. crassa* RRP-3, while the third one, including the homologous of *N. crassa* QDE-1, contains 3 *G. margarita* (g26014.t1, g25422.t1 and g15894.t1) and 3 *R. irregularis* sequences. The presence in *G. margarita* and *R. irregularis* of homologous sequences of the three well characterized *N. crassa* RdRp (Nakayashiki et al., 2006) suggests that AMF are equipped with a complete set of RdRp involved in the canonical *N. crassa* fungal RNAi pathway. Remarkably, one *G. margarita* RdRp (g15894.t1), which is distantly related to other fungal proteins, is characterized by the presence of some unusual “AAA” domains (“ATPases associated with diverse cellular activities”) at its C-terminal, which may suggest a distinct molecular function. Furthermore, we did not detect in *G. margarita* the occurrence of the small RdRp peptides found in *R. irregularis* (Silvestri et al., 2019).

Interestingly, we also found a protein of 232 amino acids (g12004.t1; not shown in the phylogenetic tree) containing an “RdRp_1” PFAM domain, which is a typical C-terminal domain of RdRp found in many eukaryotic viruses (Interpro ID: IPR001205). A blastp analysis against “non-redundant protein sequences” database on NCBI (E-value $\leq 1e-5$) revealed that g12004.t1 is similar to a number of hypovirus- and fusarivirus sequences; the best viral hit is the polyprotein of *Cryphonectria* hypovirus 4 (NCBI accession: YP_138519.1). Notably, the same blast analysis highlighted 5 protein sequences of the phylogenetically related AMF *Gigaspora rosea* (NCBI accession: RIB25480, RIB25479, RIB01634, RIB24634, RIB12248). This result provides evidence of an endogenization event of a hypovirus presumably occurring in an ancestor of the *Gigaspora* lineage. It is worth noting that no hypovirus has ever been reported in AMF.

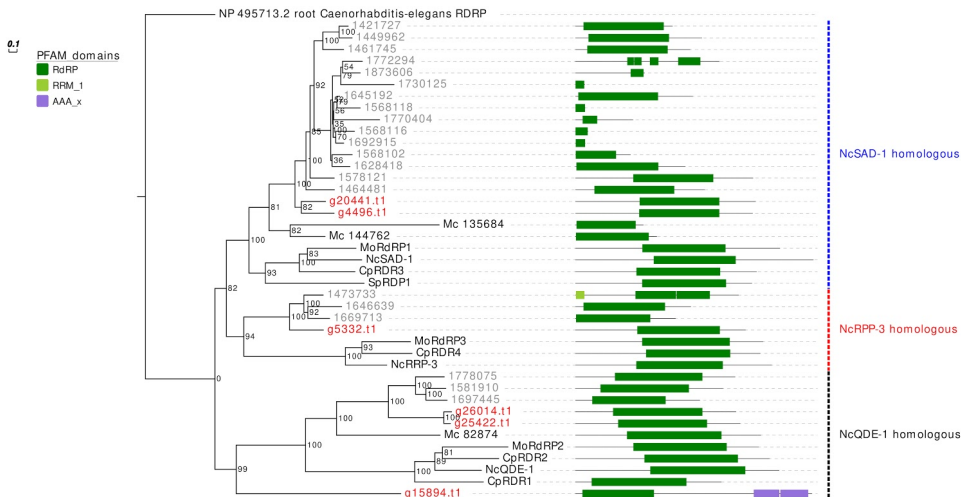


Figure 2. Phylogenetic analysis and PFAM domains of RdRp proteins. Sequences are discernible by species according to a two-letter prefix and/or a color code: Mo = *Magnaporthe oryzae*, Nc = *Neurospora crassa*, Mc = *Mucor circinelloides*, Sp = *Schizosaccharomyces pombe*, Cp = *Cryphonectria parasitica*, grey = *Rhizophagus irregularis*, red = *Gigaspora margarita*. Protein ID (NCBI or JGI): MoRdRP1 = XP_003721007.1, MoRdRP2 = XP_003711624.1, MoRdRP3 = XP_003712093.1, NcQDE-1 = EAA29811.1, NcSAD-1 = XP_964248.3, NcRRP-3 = XP_963405.1, SpRDP1 = NP_001342838.1, McRdRP-1 = 111871, McRdRP-2 = 104159, CpRDR1 = 270014, CpRDR2 = 35624, CpRDR3 = 10929, CpRDR4 = 339656. *R. irregularis* proteins are identified by JGI numeric codes. *G. margarita* proteins are identified by Venice et al. (2019) annotation code. The numbers at the nodes are bootstrap values (%) for 1000 replications. Tree was rooted using *Caenorhabditis elegans* RdRP (NCBI Reference Sequence: NP_495713.2). Tree was reshaped on root with Newick Utilities v1.6 and visualized with Evolview v3.

Concerning the DCL phylogeny, the single *G. margarita* protein sequence clusters together with the *R. irregularis* one (Figure 3). However, the presence of one DCL may not be a common AMF trait, since we previously reported 2 sequences in *R. clarus* (Silvestri et al. 2019).

We then searched for evidence of expression of the genes encoding for the RNAi-related proteins exploiting the transcriptomic data published by Venice et al. (2019) obtained from 4 different conditions: germinating spores, strigolactone-treated spores, extraradical and intraradical mycelium. All the genes are expressed in all the conditions (Supplementary Figure 1) with the only exception of the endogenized viral fragment (*g12004.t1*), which shows no

expression or very low expression levels in symbiotic (extraradical and intraradical mycelium) and in the asymbiotic (spores) conditions, respectively. In conclusion, the data confirmed that AMF are equipped with an RNAi machinery, characterized by the expansion of the *AGO-like* and, to some extent, the *RdRp* gene families. It would be interesting to understand whether these gene expansions were followed by functional differentiation, as happened to plant *AGO* (Poulsen et al., 2013).

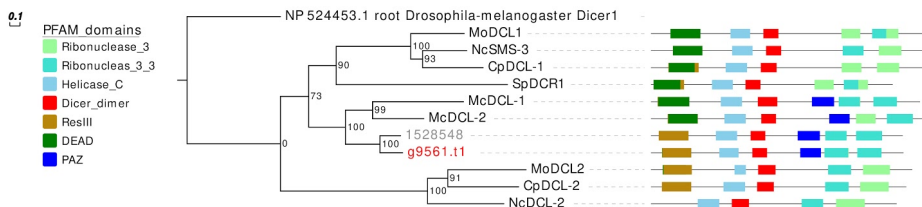


Figure 3. Phylogenetic analysis and PFAM domains of DCL proteins. Sequences are discernible by species according to a two-letter prefix and/or a color code: Mo = *Magnaporthe oryzae*, Nc = *Neurospora crassa*, Mc = *Mucor circinelloides*, Sp = *Schizosaccharomyces pombe*, Cp = *Cryphonectria parasitica*, grey = *Rhizophagus irregularis*, red = *Gigaspora margarita*. Protein ID (NCBI): MoMDL1 = XP_003714515.1, MoMDL2 = XP_003715365.1, NcSMS-3 = XP_961898.1, NcDCL-2 = XP_963538.3, SpDCR1 = NP_588215.2, McDCL-1 = CAK32533.1, McDCL-2 = CAZ65730.1, CpDCL-1 = ABB00356.1, CpDCL-2 = ABB00357.1. *R. irregularis* proteins are identified by JGI numeric codes. *G. margarita* proteins are identified by Venice et al. (2019) annotation code. The numbers at the nodes are bootstrap values (%) for 1000 replicates. Tree was rooted using *Drosophila melanogaster* Dicer 1 (NCBI Reference Sequence: NP_524453.1). Tree was reshaped on root with Newick Utilities v1.6 and visualized with Evolvew v3.

It is tempting to speculate that this uncommonly high number of *AGO-like* could be related to the large amount of transposable elements of AMF genomes (Muszewska et al., 2017). In this context, we hypothesize that specific classes of *AGO-like* may be involved in the defense against mobile elements. It is worth noting that *G. margarita* lacks some genome defense mechanisms characterized in other fungi, such as fungal repeat-induced point mutation (RIP) and meiotic silencing of unpaired DNA (MSUD) (Venice et al., 2019); furthermore, based on AMF so far sequenced, TEs invasion seems to be specific of Gigasporaceae. In this context an efficient and fine-tuned anti-transposable elements defense system based on RNAi-related pathway could be instrumental in maintaining genome integrity. Moreover, since recent

indirect evidences suggest that sRNAs are exchanged between plant and fungi in the AM symbiosis (Helber et al., 2011; Kikuchi et al., 2016; Tsuzuki et al., 2016; Xie et al., 2016; Voß et al., 2018), we speculate that the *AGO-like* gene expansion (and possibly their functional differentiation) in AMF mirrors the need to process, in a finely-tuned way, the information that may come from the host plant. Further functional analyses will be needed to validate this hypothesis.

***G. margarita* is characterized by a peculiar small RNA population**

The role of sRNAs in AMF is still largely unknown. A preliminary characterization has been so far reported only for the model species *R. irregularis* (Silvestri et al., 2019). In this context the main focus of this work was to characterize the small RNAome of *G. margarita*, a species hosting a complex viral and endobacterial population (Ghignone et al., 2012; Turina et al., 2018) which has not been found in *R. irregularis* and *G. rosea* isolates so far. We sequenced three *G. margarita* sRNA libraries, each constructed with RNA extracted from germinated spores. The Illumina platform produced a total of 73,308,493 sRNA reads which were first cleaned up from adapters, artifacts, low-quality reads; after the removal of tRNA-, rRNA-, snRNA- and snoRNA-related sequences, a total of 31,101,040 18-35-nt long reads were kept for further analysis (Supplementary Data Sheet 1).

The resulting sRNA sequences were mapped with no mismatches to the nuclear and mitochondrial genomes of *G. margarita* (Pelín et al., 2012; Venice et al., 2019), to the genome of the endobacterium *CaGg* (Ghignone et al., 2012) and to the genomes of the 6 viruses (4 mitoviruses, one Giardia-like and one Ourmia-like virus) identified in *G. margarita* (Turina et al., 2018). Sixtyone-64% of the reads, depending on the sRNA library, mapped exclusively to the nuclear genome, about 6% to the mitochondrial genome and the 0.3-0.8% to the endobacterial genome. The amount of reads uniquely mapping to the mitoviral genomes varied from 2.3-3.3% for Mitovirus 1 to 0.5-0.7% for Mitovirus 4. About 0.01% of the reads mapped to the genome of the Ourmia-like virus, while only about 0.0002% of total reads were associated with the Giardia-like virus (Figure 4). The reads mapped to both strands of viral genomes with different percentages (Supplementary Table 1). A limited number (0.13%) of total sRNA reads mapped to both nuclear

genome and mitochondrial or endobacterial or mitoviral genomes (Supplementary Figure 2).

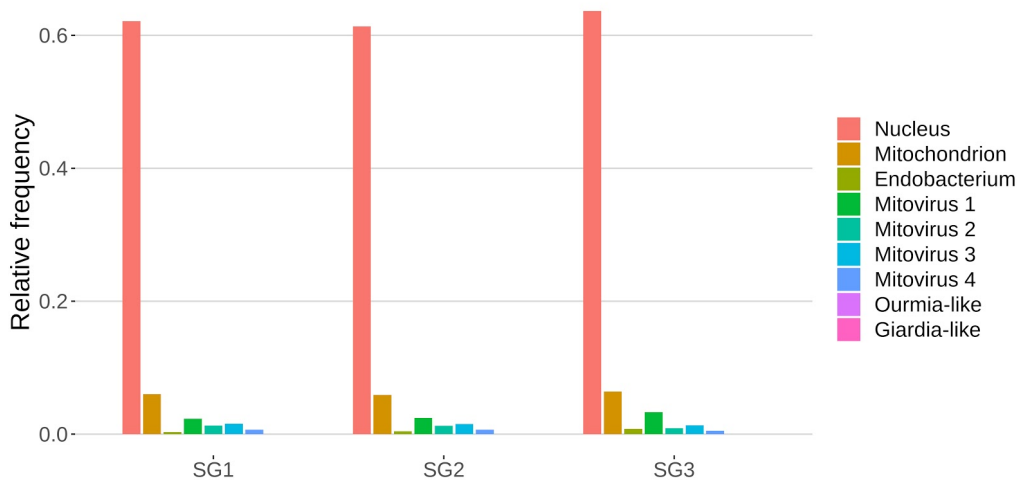


Figure 4. Relative mapping frequencies of small RNA to *Gigaspora margarita* metagenome (SG1, SG2 and SG3 refer to different libraries of “germinated spores” samples).

The redundant and non-redundant sRNA reads mapping to all the different genomes, with the exception of the endobacterial one, were characterized by a similar unimodal nucleotide length distribution in which the most representative class consisted of the 25-nt long sequences (and 24-nt as the second most abundant class for all of them) (Figure 5). The accumulation of sRNAs with specific nucleotide lengths is commonly associated with the presence of an active sRNA-generating pathway (Mueth et al., 2015), since fungal species that do not possess a functional RNAi, such as *S. cerevisiae* (Drinnenberg et al., 2009), or DCL knock-out mutants (Raman et al., 2017) are not characterized by peaks over 20 nt. The nucleotide length distribution of the reads mapping to the endobacterial genome showed no evident peak; these sRNAs are likely to have originated by random degradation of longer transcripts, in accordance with the model that does not contemplate the presence of active RNAi mechanism in prokaryotes.

The nucleotide size distribution observed for *G. margarita* nuclear DNA mapping sRNAs (*Gma*-sRNAs; unimodal with maximum at 25-nt) is different from that of *R. irregularis* (bimodal with maxima at 24- or 26-nt and 31- to 33-nt; Silvestri et al., 2019). This result is not surprising considering that the

length of the sRNAs seems not to be a conserved trait in fungi, even among species from the same genus, such as *Fusarium oxysporum* and *Fusarium graminearum* (Chen et al., 2014, 2015). For example *N. crassa* mainly produces 25-nt long sRNAs (Fulci and Macino, 2007), *M. circinnelloides* 21- and 25-nt (Nicolás et al., 2003), *Aspergillus nidulans* 25-nt (Hammond and Keller, 2005), *M. oryzae* 19- to 23-nt (Kadotani et al., 2003), *Cryptococcus neoformans* 22-nt (Dumesic et al., 2013), *Trichoderma atroviride* 20-21- and 24-nt (Carreras-Villaseñor et al., 2013), *F. graminearum* 27- and 28-nt (Chen et al., 2015), *F. oxysporum* 19- and 21- nt (Chen et al., 2014), *Sclerotinia sclerotiorum* 22-nt (Derbyshire et al., 2019) and *Puccinia striiformis* 22-nt (Mueth et al., 2015).

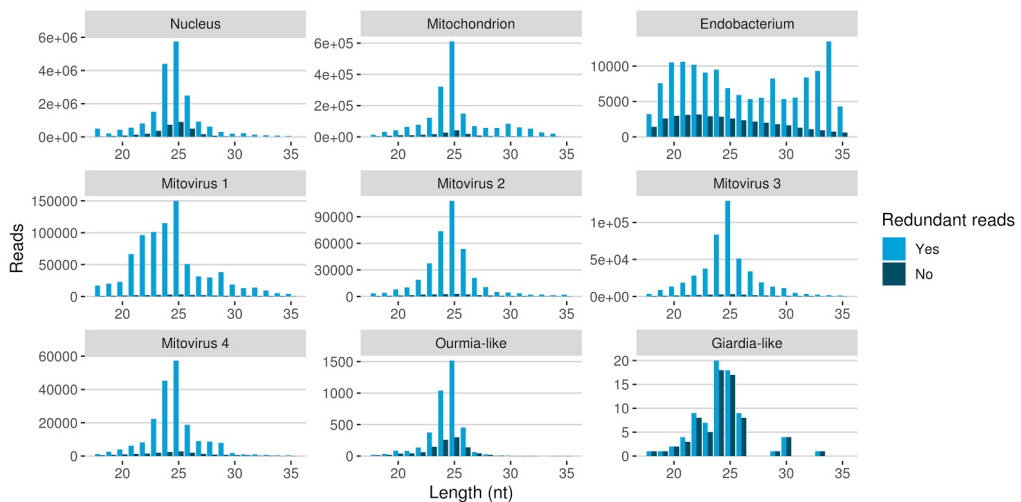


Figure 5. Nucleotide size distribution of sRNA reads (redundant and non-redundant) mapping to *G. margarita* metagenome.

The differences of the nuclear small RNAome composition between the two AMF can also be partially explained by the different experimental setups: while for *R. irregularis* we sequenced sRNAs from symbiotic conditions (fungal mycelium growing inside or outside plant roots), for *G. margarita* we sequenced sRNAs from asymbiotic condition (axenically germinated spores). A further interesting observation was the high level of sRNA reads mapping to genomes of the four mitoviruses identified in *G. margarita* (Turina et al., 2018). Mitoviruses generally replicate in their host's mitochondria (Hillman and Cai, 2013) thus in this case, exploiting the mitochondrial translation machinery,

they rely on the mitochondrial translation code (such as the UGA codon for tryptophan, which in nuclear genetic code is a stop codon). Interestingly, *G. margarita* mitoviral genetic code lacks UGA codons and can be virtually translated in both cytosol and mitochondria (Turina et al., 2018). Recently, the sRNA response to a strictly mitochondrial plant mitovirus (Chenopodium quinoa mitovirus 1) was characterized: it has been shown that the mitovirus escaped the antiviral RNAi that normally originates 21-22 nt sRNA for cytoplasmic viruses. The overall number of sRNA of mitoviral origin was very low compared to a cytoplasmic virus, and the most represented length was 17 nt (Nerva et al., 2019), corresponding to the average size of sRNA originated inside plant mitochondria. In another study the sRNA response to a mitovirus in *F. circinatum* also pointed to a protection from the cytoplasmic antiviral RNAi since, also in this case, a relatively low accumulation of sRNA had the same length distribution of mitochondrial sRNAs (Muñoz-Adalia et al., 2018). A similar situation has been observed for the mitovirus infecting the ascomycete *C. parasitica* (Shahi et al., 2019). The analysis of mitoviral sRNA in *G. margarita* did not allow us to discern whether mitoviruses replicate in mitochondria or cytosol since, contrary to what happens in plants, we do not detect differences in the nucleotide length profiles for sRNAs with nuclear and mitochondrial origins. Both are in fact characterized by 25-nt long sRNA peaks, the same of sRNAs with mitoviral origin, suggesting the presence of similar sRNA-generating processes in the two cell compartments. Notably, this seems to be a specific feature of *G. margarita*; the analysis of nucleotide length profile of *R. irregularis* sRNAs (exploiting sRNA-seq data previously published; Silvestri et al., 2019) revealed that the population of mitochondrial DNA mapping sRNAs (decreasing curve from 18- to 35-nt) clearly differs from the population of nuclear DNA mapping sRNAs (Supplementary Figure 3).

The very high number of sRNA accumulating during mitovirus infection seem to suggest that, contrary to some of the systems described above, their RNA is indeed targeted by cytoplasmic RNAi, possibly during promiscuous replication (both mitochondrial and cytoplasmic). If this is a true antiviral response remains to be established. Another characteristic of *G. margarita* mitoviruses is the production during replication of episomic DNA fragments corresponding to their sequence (Turina et al., 2018): our sRNA analysis also aimed at searching for evidence of a specific anti-viral response originated by such DNA fragments, in analogy to the PIWI sRNA response originated by

DNA fragments to control RNA viruses in insects (Goic et al., 2016). However, we could not detect any peculiar population of viral-derived sRNA (such as the insect PIWI sRNA) that allowed us to envisage a similar role for *G. margarita* viral DNA fragments.

Mapped sRNA reads were also analyzed for their 5'-end nucleotide composition (normalized on nucleotide composition of each genome). A general enrichment in uracil was observed for 23-26-nt long sequences with few differences depending on their genomic origin (Figure 6). Interestingly, the 23-26 nt range also corresponded to the group of the most expressed sRNAs (Figure 5), with the only exception of the endobacterium and the Giardia-like virus, the latter however characterized by very few total mapped reads. The 5'-ends enrichment in uracil, which is a rather common feature of fungal sRNA (Nicolás et al., 2003; Dumesic et al., 2013; Mueth et al., 2015; Nguyen et al., 2018; Derbyshire et al., 2019; Silvestri et al., 2019), for the 23-26-nt long *G. margarita* sRNAs (the most expressed ones), could be a further indication of a functional role. Remarkably, the nucleotide composition of the 5'-end of sRNAs affects their ability to be loaded onto different classes of AGO proteins; in *A. thaliana* the uracil at the 5'-end is indeed associated with the sRNA loading onto AGO1 and AGO10, while cytosine is associated with AGO5, and adenine with AGO2, AGO4, AGO6, AGO7 and AGO9 (Borges and Martienssen, 2015). Furthermore, it is worth noting that the 5' end nucleotide compositions of the 23-26 nt range mitoviral sRNAs is enriched in uracil. This is a common feature observed for several mycoviruses (Donaire and Ayllón, 2017).

All these findings suggest the presence of active molecular pathways producing sRNAs, with different genomic origins (nuclear, mitochondrial or viral). The sRNA peaks at 25 nt for all the reads mapping to different genomes (with the exception of the endobacterial one), together with the uracil enrichment at their 5'-end, suggests the presence of an RNAi pathway in *G. margarita* able to process sRNAs with nuclear and mitochondrial origin and also able to specifically target viral sequences.

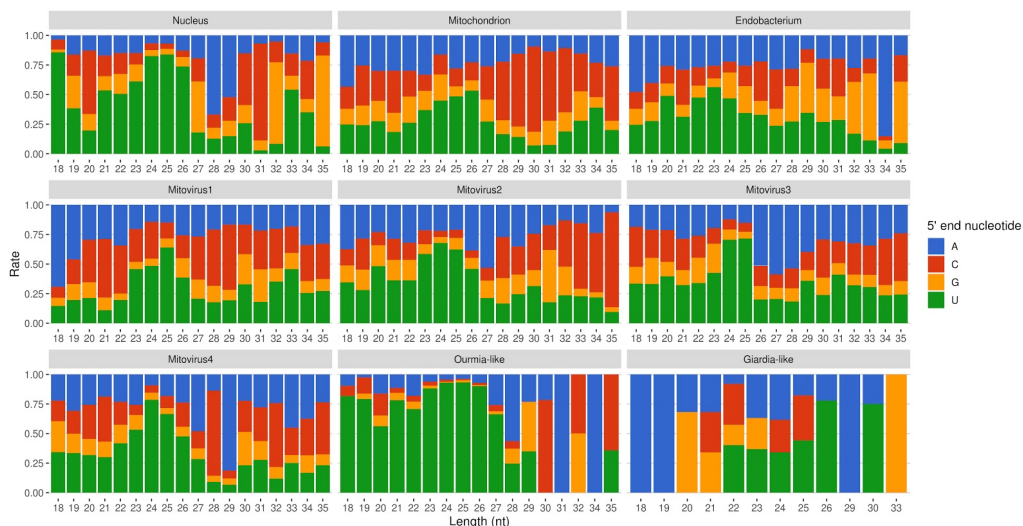


Figure 6. Relative nucleotide frequency of the 5' ends of the sRNAs reads mapping to *Gigaspora margarita* metagenome (normalized on nucleotide composition of each genome)

***G. margarita* is characterized by different populations of nuclear sRNA-generating loci**

Using ShortStack (Johnson et al., 2016) we performed a prediction and annotation of *G. margarita* nuclear genome loci producing sRNAs (*Gma*-sRNA-generating loci). Applying the same parameters used in a previous study (Silvestri et al., 2019), we predicted 4575 loci of which 3422 (75%) localized in intergenic regions and 1153 (25%) overlapped with predicted protein-encoded genes (i.e. loci that shared, for at least one nucleotide, the same genomic positions of elements reported as “mRNA” in the annotation file; Supplementary Data Sheet 2). A different situation was reported in *R. irregularis* that showed a predominance of sRNA-generating loci overlapping with protein-encoding genes (67% of the total; Silvestri et al., 2019). This could be due to a higher occurrence of intergenic regions in the genome of *G. margarita* considering that its very large genome contains a number of predicted genes similar to *R. irregularis* (Venice et al., 2019).

A total of 762 (17%) *Gma*-sRNA-generating loci showed similarity to fungal repetitive elements from RepBase 23.04 (Supplementary Data Sheet 2), in analogy to the 11% observed in *R. irregularis* (Silvestri et al., 2019). A further analysis of the genomic positions of the 4575 *Gma*-sRNA-generating loci revealed that 3635 (79%) overlapped for at least one nucleotide with *G.*

margarita transposable elements (Supplementary Data Sheet 2); 672 of these loci overlapped on both transposable elements and protein-encoding genes. A PCA analysis (Figure 7) on nuclear sRNA-generating loci was then performed, as previously proposed (Fahlgren et al., 2013), taking into account 19 independent variables per locus (locus length in nucleotide, total number of mapped reads and their nucleotide size proportion from 18- to 35-nt). The 23.4% and 9.4% of the total variance were explained by PC1 (principal component 1) and PC2, respectively. Six variables (the proportion of 30-, 31-, 29-, 28- and 25-nt long sequences) contribute for more than 50% to PC1 (Figure 7A), which mainly separates the loci in two groups, as confirmed by HDBSCAN (density-based spatial clustering of applications with noise) algorithm (Campello et al., 2015) (Figure 7B). The two clusters, “cluster 1” and “cluster 2”, were composed by 172 and 4180 loci, respectively, while the remaining 223 loci were not clustered (we renamed this group as “cluster 0”; Supplementary Data Sheet 2). The nucleotide size distributions of the sRNA reads mapping to the loci belonging to the three clusters showed different average profiles. In particular, we observed an unimodal curve with the maximum peak at 25 nt for cluster 2, and a flat curve for cluster 0 and cluster 1, the latter however slightly enriched in 22-26 nt long sequences (Figure 8). The loci belonging to cluster 2 could be further differentiated in different groups based on their maximum peaks. In fact, while 25 nt was the most represented size for the majority of them (3091 loci), the remaining were characterized by maximum peaks at 23 nt (31 loci), 24 nt (901 loci), 26 nt (113 loci), 27 nt (2 loci) or at both 24 and 25 nt (22 loci) (Supplementary Figure 4). The presence of loci producing sRNAs of different length suggests the existence of at least partially different sRNA-generating processes, specifically acting on each group of sRNA-generating loci from “cluster 2”. Unfortunately, the lack of stable genetic transformation protocols for AMF, and so the possibility to obtain DCL knock-out mutants, makes it difficult to understand whether all these non-25-nt-sRNA-generating loci are dependent, for sRNA production, on the single *G. margarita* DCL or whether other DCL-independent processes are involved, as reported in the basal fungus *M. circinelloides* (Torres-Martínez and Ruiz-Vázquez, 2016). The setup of complementation assays, in which *G. margarita* DCL is expressed in *M. circinelloides* DCL knock-out mutants, could help understanding whether the *G. margarita* single DCL participate in the biogenesis of the non-25-nt long sRNAs.

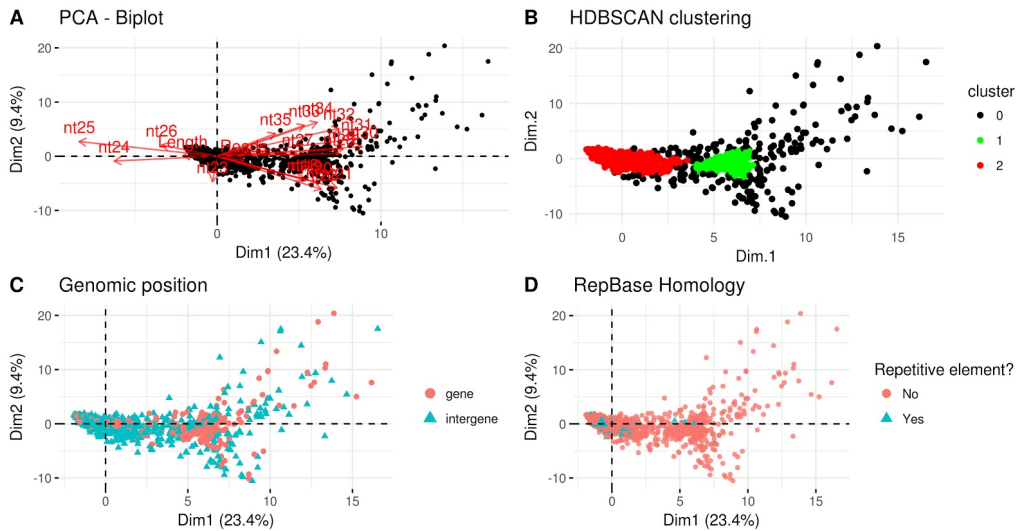


Figure 7. Characterization of *Gigaspora margarita*(*Gma*)-sRNA generating loci. **(A)** Biplot of principal component 1 and 2 of PCA based on the length of loci, the total number of mapped reads and the nucleotide size proportion of *Gma*-sRNAs (from 18 nt to 35 nt) defining each locus (19 total variables). **(B)** HDBSCAN clustering reveals the presence of two distinct populations of data (Cluster 1 and 2). **(C)** Overview of the relative genomic positions of the loci compared to those of protein-encoding genes. **(D)** Overview of the homology of *Gma*-sRNA-generating loci with fungal repetitive elements in RepBase 23.04 (tblastx: E-value ≤ 0.00005).

The loci belonging to different clusters also differentiated on the base of their relative genomic positions, with an amount of intergenic loci of about 62% and 78% for cluster 0 and cluster 2 respectively, as opposed to the 23% from cluster 1 (Figure 7c). A further differentiation was evident analyzing the percentage of repetitive elements homologous, with only 2% and 1% of the total loci from cluster 0 and cluster 1, opposed to the 18% from cluster 2 (Figure 7d).

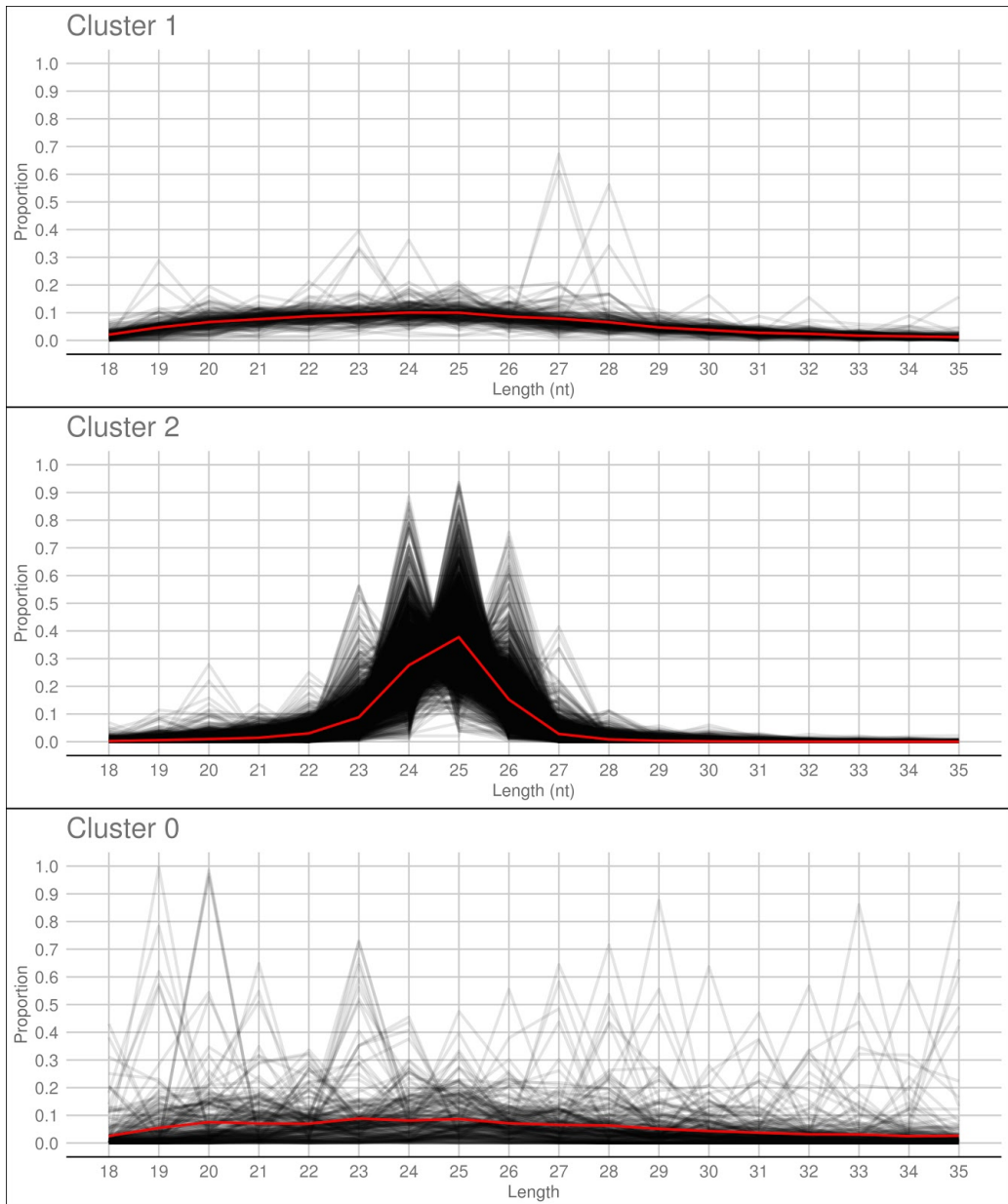


Figure 8. Nucleotide size distribution of sRNA reads that define the *Gigaspora margarita*-sRNA-generating loci of Cluster 1, Cluster 2 and Cluster 0 (the latter containing the non-clustered loci) according to HDBSCAN clustering. Black lines refer to the nucleotide size distribution of the sRNA reads defining the individual loci and red lines to the average nucleotide size distribution of each cluster.

The comparison with *R. irregularis* revealed that the sRNA-generating loci belonging to “Cluster 2” in both species share several features, such as the relative amount of intergenic loci (63% vs 78%; *R. irregularis* vs *G. margarita*) and repetitive elements homologs (16% vs 18%) and the average length of mapped sRNAs (unimodal curve with maximum at 24-nt vs unimodal curve with maximum at 25-nt). Instead, for “Cluster 1”, the differences between the two species are more relevant, relative to both percentage of total loci belonging to cluster (52% vs 4%) and average length of mapped sRNAs (decreasing curve from 18- to 35-nt with no evident peaks vs flat curve slightly enriched in 22- to 26-nt long sequences).

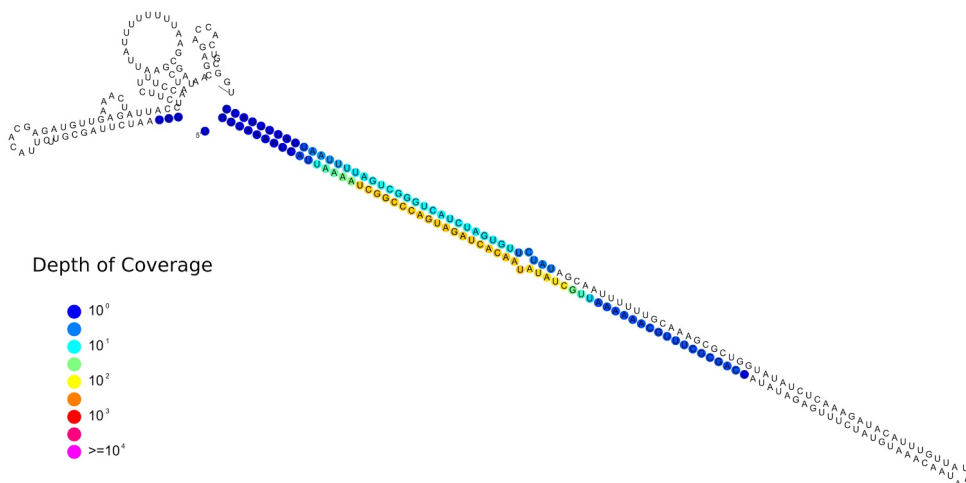


Figure 9. Predicted secondary structures of the putative *Gigaspora margarita* miRNA-like (Locus 1809; scaffold_534:156601-156872, positive strand) with color-coded sRNA-seq coverage per nucleotide.

A blastn analysis revealed that 261 *Gma*-sRNA-generating loci are highly similar to 132 *R. irregularis*-sRNA-generating loci (evalue $\leq 1e-5$); 135 out of 261 are intergenic loci (of which 18 correspond to transposable elements), while the remaining ones overlap with annotated genes (Supplementary Data Sheet 2). The presence of different populations of nuclear loci producing sRNAs may indicate the occurrence of different sRNA-generating processes.

One sRNA-generating locus (Locus 1809) was predicted as miRNA-like by ShortStack (Figure 9). Together with the 10 previously annotated in *R. irregularis*, this is the first evidence of the presence miRNA-like loci in basal fungi as no miRNA-like sequence was described in *M. circinelloides* (Torres-Martínez and Ruiz-Vázquez 2017).

Finally, we observed that the genomic coordinates of the gene encoding for the endogenized viral RdRp fragment (*g12004.t1*; scaffold_323, bp from 157181 to 157959, positive strand) is totally contained inside the genomic boundaries of the sRNA-generating locus 1266 (scaffold_323; bp from 155733 to 160403) but on the opposite strand (since about the 98% of locus 1266 sRNAs mapped to the negative strand with a coverage of about 317 reads per million; Supplementary Data Sheet 2). Considering that the RNA-seq data from germinated spores pointed to a very low expression level for *g12004.t1* (Supplementary Figure 1), we can hypothesize that the occurrence of abundant antisense sRNAs is evidence of a silencing activity towards the endogenized virus fragment.

***G. margarita* fungal/viral sRNA can potentially target plant transcripts**

Cross-kingdom RNAi is nowadays recognized as a key mechanism involved in several plant-microbe interactions (Huang et al., 2019). Some experiments, which applied HIGS and VIGS techniques on mycorrhizal plants, suggested that sRNAs can also be exchanged between plants and AMF (Helber et al., 2011; Kikuchi et al., 2016; Tsuzuki et al., 2016; Voß et al., 2018), including *G. margarita* (Xie et al., 2016). In this context, we performed an *in silico* analysis in order to identify *G. margarita* sRNAs potentially able to silence plant transcripts, using *Medicago truncatula* as model organism (Bell, 2001). Our work is based on the assumption that some sRNAs from germinated spores are also maintained in the symbiotic phase; we focused therefore only on the most expressed (> 100 RPM) 21-nt long *Gma*-sRNAs. The 21-nt long sRNAs are indeed generally involved in plants post-transcriptional gene silencing and, in addition, they are generally loaded onto AGO1 (Axtell, 2013), the only class of AGO known so far to be involved in cross-kingdom RNAi (Weiberg et al. 2013; Wang et al. 2016; Ren et al. 2019; Cui et al. 2019). Following this approach, we identified 292 *M. truncatula* mRNAs potentially targeted by 27 *Gma*-sRNAs (Supplementary Data Sheet 3), enriched in 3 GO terms: “signal transducer activity” (GO:0004871), “molecular transducer activity”

(GO:0060089) and “lipid binding” (GO:0008289), according to Plant GO Slim ontology. A similar analysis, conducted on the most abundant (> 100 RPM) 21-nt long sRNAs of the AMF *R. irregularis* (*Rir*-sRNAs; Silvestri et al. 2019), revealed 663 *M. truncatula* transcripts as putative targets of 57 *Rir*-sRNAs (Supplementary Data Sheet 4). Eleven plant mRNA sequences were predicted as targets of both *Gma*-sRNAs and *Rir*-sRNAs (Supplementary Table 2); notably, three of them - a “chitinase” (AES62408), an “expansin A10” (AES77475) and a “CCR4-NOT transcription complex protein” (AET01158) encoding genes - are targeted at the same site by sRNAs of both AMF (Supplementary Data Sheet 3 and 4), reinforcing the hypothesis of a possible involvement of these sRNAs in host gene regulation in the AM symbiosis.

Since we assumed a possible high rate of false positives of these *in silico* prediction, as a negative control, we also performed a further target prediction on *M. truncatula* transcriptome using sRNAs from the unrelated non mycorrhizal fungus *Aspergillus fumigatus* (Özkan et al., 2017). We identified 309 plant mRNA targets of 43 abundant (> 100 RPM) *A. fumigatus* sRNAs (*Afu*-sRNAs; Supplementary Data Sheet 5). Among these, 6 plant mRNAs were predicted as targets of both *Gma*-sRNAs and *Afu*-sRNAs, six of both *Rir*-sRNAs and *Afu*-sRNAs and one (KEH44371, an “Ubiquitin-conjugating enzyme E2”) of sRNAs from all three fungal species (Supplementary Table 2). These results confirm that *in silico* target prediction analyses must only be intended as a preliminary step for the identification of target genes, especially in plant-microbe cross-kingdom interactions. *In vivo* experiments would be necessary to demonstrate the AMF sRNA silencing effect against *in silico* identified target *M. truncatula* sequences (Zanini et al., 2018). Our results also suggest that additional sRNA data from a higher number of AMF species will be instrumental to increase the robustness of *in silico* target prediction to identify conserved AMF sRNA effectors as well as plant targets.

We also investigated the potential involvement of *G. margarita* viral sRNAs in the regulation of plant mRNAs through cross-kingdom RNAi: we detected, by the *in silico* analysis, 248 plant transcripts potentially targeted by the 55 most expressed (> 100 RPM) 21-nt long sRNAs derived from *G. margarita* mitoviruses (Supplementary Data Sheet 6). Although the role of viral-derived sRNAs in host plant gene regulation has been extensively studied and characterized (Shimura et al., 2011; Smith et al., 2011; Adkar-Purushothama et

al., 2015; Yang et al., 2019), this is a first evidence of their potential implication in cross-kingdom RNAi from a fungus to a plant.

In conclusion, our work indicates that several AM fungal/viral sRNAs could potentially target plant transcripts and could act, if they are transferred into plant host cells, as RNA effectors in the AM symbiosis. Further experiments are needed to validate these putative fungal/viral sRNA - plant mRNA target pairs.

5.5. Conclusion

Our work demonstrates that AMF *G. margarita* is equipped with a complete RNAi machinery and, in analogy to *Rhizophagus* species, shows a peculiar expansion of the *AGO* gene family. We also provided the first characterization of the small RNAome of a complex symbiotic meta-genome that involves a fungus permanently associated to bacteria and viruses. Furthermore, our results point to a possible cross-kingdom interaction with the plant host mediated by some of these sRNAs, including those of viral origin.

Mining additional AMF genomes and small RNAome will provide new insights on how these ancient microbes are able to establish long-lasting interactions with plants, bacteria and viruses.

Data Availability

The sRNA-seq datasets analyzed in this study are available on European Nucleotide Archive (ENA) under the study accession number PRJEB35457.

Author Contributions

LL, AS and MT designed experiments; AS, VF, LM and FV performed experiments and, together with PB, MT and LL, carried out data analyses. AS and LL wrote the manuscript.

Funding

The research was supported by the Project ex-post 2018 from the University of Turin - Compagnia di San Paolo.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

The research was supported by the Project ex-post 2018 from the University of Turin - Compagnia di San Paolo.

5.6. References

- Adkar-Purushothama, C. R., Brosseau, C., Giguère, T., Sano, T., Moffett, P., and Perreault, J. P. (2015). Small RNA derived from the virulence modulating region of the potato spindle tuber viroid silences callose synthase genes of tomato plants. *Plant Cell* 27, 2178–2194. doi:10.1105/tpc.15.00523.
- Axtell, M. J. (2013). Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol.* 64, 137–159. doi:10.1146/annurev-arplant-050312-120043.
- Bell, C. J. (2001). The Medicago Genome Initiative: a model legume database. *Nucleic Acids Res.* 29, 114–117. doi:10.1093/nar/29.1.114.
- Borges, F., and Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nat. Publ. Gr.* 16, 1–15. doi:10.1038/nrm4085.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421. doi:10.1186/1471-2105-10-421.
- Campello, R. J. G. B., Moulavi, D., Zimek, A., and Sander, J. (2015). Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Trans. Knowl. Discov. Data* 10, 1–51. doi:10.1145/2733381.
- Carreras-Villaseñor, N., Esquivel-Naranjo, E. U., Villalobos-Escobedo, J. M., Abreu-Goodger, C., and Herrera-Estrella, A. (2013). The RNAi machinery regulates growth and development in the filamentous fungus *Trichoderma atroviride*. *Mol. Microbiol.* 89, 96–112. doi:10.1111/mmi.12261.
- Chang, S.-S., Zhang, Z., and Liu, Y. (2012). RNA interference pathways in fungi: mechanisms and functions. *Annu. Rev. Microbiol.* 66, 305–323. doi:10.1146/annurev-micro-092611-150138.
- Chen, E. C. H., Morin, E., Beaudet, D., Noel, J., Yildirim, G., Ndikumana, S., et al. (2018). High intraspecific genome diversity in the model arbuscular mycorrhizal symbiont *Rhizophagus irregularis*. *New Phytol.* doi:10.1111/nph.14989.
- Chen, R., Jiang, N., Jiang, Q., Sun, X., Wang, Y., Zhang, H., et al. (2014). Exploring microRNA-like small RNAs in the filamentous fungus *Fusarium oxysporum*. *PLoS One* 9. doi:10.1371/journal.pone.0104956.
- Chen, Y., Gao, Q., Huang, M., Liu, Y., Liu, Z., Liu, X., et al. (2015). Characterization of RNA silencing components in the plant pathogenic fungus *Fusarium graminearum*. *Sci. Rep.* 5, 12500. doi:10.1038/srep12500.
- Chow, F. W. N., Koutsovoulos, G., Ovando-Vázquez, C., Neophytou, K., Bermúdez-Barrientos, J. R., Laetsch, D. R., et al. (2019). Secretion of an Argonaute protein by a parasitic nematode and the evolution of its siRNA guides. *Nucleic Acids Res.* 47, 3594–3606. doi:10.1093/nar/gkz142.

- Cui, C., Wang, Y., Liu, J., Zhao, J., Sun, P., and Wang, S. (2019). A fungal pathogen deploys a small silencing RNA that attenuates mosquito immunity and facilitates infection. *Nat. Commun.* 10, 4298. doi:10.1038/s41467-019-12323-1.
- Dai, X., Zhuang, Z., and Zhao, P. X. (2018). psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* 46, W49–W54. doi:10.1093/nar/gky316.
- Derbyshire, M., Mbengue, M., Barascud, M., Navaud, O., and Raffaele, S. (2019). Small RNAs from the plant pathogenic fungus *Sclerotinia sclerotiorum* highlight host candidate genes associated with quantitative disease resistance. *Mol. Plant Pathol.* 20, 1279–1297. doi:10.1111/mpp.12841.
- Donaire, L., and Ayllón, M. A. (2017). Deep sequencing of mycovirus-derived small RNAs from *Botrytis* species. *Mol. Plant Pathol.* 18, 1127–1137. doi:10.1111/mpp.12466.
- Drinnenberg, I. A., Weinberg, D. E., Xie, K. T., Mower, J. P., Wolfe, K. H., Fink, G. R., et al. (2009). RNAi in budding yeast. *Science* (80-.). 326, 544–550. doi:10.1126/science.1176945.
- Dumesic, P. A., Natarajan, P., Chen, C., Drinnenberg, I. A., Schiller, B. J., Thompson, J., et al. (2013). Stalled spliceosomes are a signal for RNAi-mediated genome defense. *Cell* 152, 957–968. doi:10.1016/j.cell.2013.01.046.
- Fahlgren, N., Bollmann, S. R., Kasschau, K. D., Cuperus, J. T., Press, C. M., Sullivan, C. M., et al. (2013). *Phytophthora* have distinct endogenous small RNA populations that include short interfering and microRNAs. *PLoS One* 8, e77181. doi:10.1371/journal.pone.0077181.
- Fulci, V., and Macino, G. (2007). Quelling: post-transcriptional gene silencing guided by small RNAs in *Neurospora crassa*. *Curr. Opin. Microbiol.* 10, 199–203. doi:10.1016/j.mib.2007.03.016.
- Ghignone, S., Salvioli, A., Anca, I., Lumini, E., Ortu, G., Petiti, L., et al. (2012). The genome of the obligate endobacterium of an AM fungus reveals an interphylum network of nutritional interactions. *ISME J.* 6, 136–145. doi:10.1038/ismej.2011.110.
- Goic, B., Stapleford, K. A., Frangeul, L., Doucet, A. J., Gausson, V., Blanc, H., et al. (2016). Virus-derived DNA drives mosquito vector tolerance to arboviral infection. *Nat. Commun.* 7, 12410. doi:10.1038/ncomms12410.
- Hahsler M, Piekenbrock M, D. D. (2019). dbSCAN: Fast Density-Based Clustering with R. *J. Stat. Softw.* 91, 1–30. Available at: doi: 10.18637/jss.v091.i01.
- Hammond, T. M., and Keller, N. P. (2005). RNA silencing in *Aspergillus nidulans* is independent of RNA-dependent RNA polymerases. *Genetics* 169, 607–617. doi:10.1534/genetics.104.035964.
- Helber, N., Wippel, K., Sauer, N., Schaarschmidt, S., Hause, B., and Requena, N. (2011). A versatile monosaccharide transporter that operates in the arbuscular mycorrhizal fungus *Glomus* sp is crucial for the symbiotic relationship with plants. *Plant Cell* 23, 3812–3823. doi:10.1105/tpc.111.089813.
- Hillman, B. I., and Cai, G. (2013). *The Family Narnaviridae. Simplest of RNA Viruses*. 1st ed. Copyright © 2013, Elsevier Inc. All Rights Reserved. doi:10.1016/B978-0-12-394315-6.00006-4.

- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. doi:10.1093/molbev/msx281.
- Huang, C.-Y., Wang, H., Hu, P., Hamby, R., and Jin, H. (2019). Small RNAs – big players in plant-microbe interactions. *Cell Host Microbe* 26, 173–182. doi:10.1016/j.chom.2019.07.021.
- Ipsaro, J. J., and Joshua-Tor, L. (2015). From guide to target: molecular insights into eukaryotic RNA-interference machinery. *Nat. Struct. Mol. Biol.* 22, 20–28. doi:10.1038/nsmb.2931.
- Johnson, N. R., Yeoh, J. M., Coruh, C., and Axtell, M. J. (2016). Improved placement of multi-mapping small RNAs. *G3: Genes, Genomes, Genetics.* 6, 2103–11. doi:10.1534/g3.116.030452.
- Junier, T., and Zdobnov, E. M. (2010). The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26, 1669–1670. doi:10.1093/bioinformatics/btq243.
- Kadotani, N., Nakayashiki, H., Tosa, Y., and Mayama, S. (2003). RNA silencing in the phytopathogenic fungus *Magnaporthe oryzae*. *Mol. Plant-Microbe Interact.* 16, 769–776. doi:10.1094/MPMI.2003.16.9.769.
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–591. doi:10.1038/nmeth.4285.
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–80. doi:10.1093/molbev/mst010.
- Kikuchi, Y., Hijikata, N., Ohtomo, R., Handa, Y., Kawaguchi, M., Saito, K., et al. (2016). Aquaporin-mediated long-distance polyphosphate translocation directed towards the host in arbuscular mycorrhizal symbiosis: application of virus-induced gene silencing. *New Phytol.* 211, 1202–1208. doi:10.1111/nph.14016.
- Krüger, M., Krüger, C., Walker, C., Stockinger, H., and Schüßler, A. (2012). Phylogenetic reference data for systematics and phylotaxonomy of arbuscular mycorrhizal fungi from phylum to species level. *New Phytol.* 193, 970–984. doi:10.1111/j.1469-8137.2011.03962.x.
- Lanfranco, L., Fiorilli, V., and Gutjahr, C. (2018). Partner communication and role of nutrients in the arbuscular mycorrhizal symbiosis. *New Phytol.* 220, 1031–1046. doi:10.1111/nph.15230.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. (2009). 2C- Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25. doi:10.1186/gb-2009-10-3-r25.
- Lê, S., Josse, J., Software, F. H.-J. of statistical, and 2008, U. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* Available at: http://factominer.free.fr/more/article_FactoMineR.pdf.
- Lee, S. J., Kong, M., Harrison, P., and Hijri, M. (2018). Conserved proteins of the RNA interference system in the arbuscularmycorrhizal fungus *rhizoglyphus irregulare* provide

- new insight into the evolutionary history of glomeromycota. *Genome Biol. Evol.* 10, 328–343. doi:10.1093/gbe/evy002.
- Mayoral, J. G., Hussain, M., Albert Joubert, D., Iturbe-Ormaetxe, I., O'Neill, S. L., and Asgari, S. (2014). *Wolbachia* small noncoding RNAs and their role in cross-kingdom communications. *Proc. Natl. Acad. Sci. U. S. A.* 111, 18721–18726. doi:10.1073/pnas.1420131112.
- Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. *Nature* 457, 413–420. doi:10.1038/nature07756.
- Mueth, N. A., Ramachandran, S. R., and Hulbert, S. H. (2015). Small RNAs from the wheat stripe rust fungus (*Puccinia striiformis* f.sp. tritici). *BMC Genomics* 16, 718. doi:10.1186/s12864-015-1895-4.
- Muñoz-Adalia, E. J., Diez, J. J., Fernández, M. M., Hantula, J., and Vainio, E. J. (2018). Characterization of small RNAs originating from mitoviruses infecting the conifer pathogen *Fusarium circinatum*. *Arch. Virol.* 163, 1009–1018. doi:10.1007/s00705-018-3712-2.
- Muszewska, A., Steczkiewicz, K., Stepniewska-Dziubinska, M., and Ginalski, K. (2017). Cut-and-Paste transposons in fungi with diverse lifestyles. *Genome Biol. Evol.* 9, 3463–3477. doi:10.1093/gbe/evx261.
- Nakayashiki, H., Kadotani, N., and Mayama, S. (2006). Evolution and diversification of RNA silencing proteins in fungi. *J. Mol. Evol.* 63, 127–135. doi:10.1007/s00239-005-0257-2.
- Nawrocki, E. P., Burge, S. W., Bateman, A., Daub, J., Eberhardt, R. Y., Eddy, S. R., et al. (2015). Rfam 12.0: Updates to the RNA families database. *Nucleic Acids Res.* 43, D130–D137. doi:10.1093/nar/gku1063.
- Nerva, L., Viganì, G., Di Silvestre, D., Ciuffo, M., Forgia, M., Chitarra, W., et al. (2019). Biological and molecular characterization of *Chenopodium quinoa* mitovirus 1 reveals a distinct sRNA response compared to cytoplasmic RNA viruses. *J. Virol.* 93, 1–17. doi:10.1128/jvi.01998-18.
- Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi:10.1093/molbev/msu300.
- Nguyen, Q., Iritani, A., Ohkita, S., Vu, B. V., Yokoya, K., Matsubara, A., et al. (2018). A fungal Argonaute interferes with RNA interference. *Nucleic Acids Res.*, 1–14. doi:10.1093/nar/gkx1301.
- Nicolás, F. E., Torres-Martínez, S., and Ruiz-Vázquez, R. M. (2003). Two classes of small antisense RNAs in fungal RNA silencing triggered by non-integrative transgenes. *EMBO J.* 22, 3983–3991. doi:10.1093/emboj/cdg384.
- Özkan, S., Mohorianu, I., Xu, P., Dalmay, T., and Coutts, R. H. A. (2017). Profile and functional analysis of small RNAs derived from *Aspergillus fumigatus* infected with double-stranded RNA mycoviruses. *BMC Genomics* 18, 1–13. doi:10.1186/s12864-017-3773-8.
- Pelin, A., Pombert, J. F., Salvioli, A., Bonen, L., Bonfante, P., and Corradi, N. (2012). The mitochondrial genome of the arbuscular mycorrhizal fungus *Gigaspora margarita* reveals

- two unsuspected trans-splicing events of group I introns. *New Phytol.* 194, 836–845. doi:10.1111/j.1469-8137.2012.04072.x.
- Poulsen, C., Vaucheret, H., and Brodersen, P. (2013). Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. *Plant Cell* 25, 22–37. doi:10.1105/tpc.112.105643.
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi:10.1093/bioinformatics/btq033.
- Raman, V., Simon, S. A., Demirci, F., Nakano, M., Meyers, B. C., and Donofrio, N. M. (2017). Small RNA functions are required for growth and development of *Magnaporthe oryzae*. *Mol. Plant-Microbe Interact.* 30, 517–530. doi:10.1094/MPMI-11-16-0236-R.
- Ren, B., Wang, X., Duan, J., and Ma, J. (2019). Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation. *Science* (80-.). 365, 919–922. doi:10.1126/science.aav8907.
- Shahi, S., Eusebio-Cope, A., Kondo, H., Hillman, B. I., and Suzuki, N. (2019). Investigation of host range of and host defense against a mitochondrially replicating mitovirus. *J. Virol.* 93, 1–15. doi:10.1128/jvi.01503-18.
- Shahid, S., Kim, G., Johnson, N. R., Wafula, E., Wang, F., Coruh, C., et al. (2018). MicroRNAs from the parasitic plant *Cuscuta campestris* target host messenger RNAs. *Nature* 553, 82–85. doi:10.1038/nature25027.
- Shimura, H., Pantaleo, V., Ishihara, T., Myojo, N., Inaba, J. ichi, Sueda, K., et al. (2011). A viral satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in chlorophyll biosynthesis using the RNA silencing machinery. *PLoS Pathog.* 7, 1–12. doi:10.1371/journal.ppat.1002021.
- Silvestri, A., Fiorilli, V., Miozzi, L., Accotto, G. P., Turina, M., and Lanfranco, L. (2019). In silico analysis of fungal small RNA accumulation reveals putative plant mRNA targets in the symbiosis between an arbuscular mycorrhizal fungus and its host plant. *BMC Genomics* 20, 169. doi:10.1186/s12864-019-5561-0.
- Smith, N. A., Eamens, A. L., and Wang, M. B. (2011). Viral small interfering RNAs target host genes to mediate disease symptoms in plants. *PLoS Pathog.* 7, 1–9. doi:10.1371/journal.ppat.1002022.
- Spatafora, J. W., Chang, Y., Benny, G. L., Lazarus, K., Smith, M. E., Berbee, M. L., et al. (2016). A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* 108, 1028–1046. doi:10.3852/16-042.
- Subramanian, B., Gao, S., Lercher, M. J., Hu, S., and Chen, W.-H. (2019). Evolveview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Res.* 47, W270–W275. doi:10.1093/nar/gkz357.
- Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., et al. (2017). AgriGO v2.0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 45, W122–W129. doi:10.1093/nar/gkx382.
- Torres-Martínez, S., and Ruiz-Vázquez, R. M. (2016). RNAi pathways in *Mucor*: A tale of proteins, small RNAs and functional diversity. *Fungal Genet. Biol.* 90, 44–52. doi:10.1016/j.fgb.2015.11.006.

- Tsuzuki, S., Handa, Y., Takeda, N., and Kawaguchi, M. (2016). Strigolactone-induced putative secreted protein 1 is required for the establishment of symbiosis by the arbuscular mycorrhizal fungus *Rhizophagus irregularis*. *Mol. Plant-Microbe Interact.* 29, 1–59. doi:10.1094/MPMI-10-15-0234-R.
- Turina, M., Ghignone, S., Astolfi, N., Silvestri, A., Bonfante, P., and Lanfranco, L. (2018). The virome of the arbuscular mycorrhizal fungus *Gigaspora margarita* reveals the first report of DNA fragments corresponding to replicating non-retroviral RNA viruses in Fungi. *Environ. Microbiol.* 00. doi:10.1111/1462-2920.14060.
- Venice, F., Ghignone, S., Salvioli, A., Amselem, J., Novero, M., Xianan, X., et al. (2019). At the nexus of three kingdoms: the genome of the mycorrhizal fungus *Gigaspora margarita* provides insights into plant , endobacterial and fungal interactions. 00. doi:10.1111/1462-2920.14827.
- Voß, S., Betz, R., Heidt, S., Corradi, N., and Requena, N. (2018). RiCRN1, a crinkler effector from the arbuscular mycorrhizal fungus *rhizophagus irregularis*, functions in arbuscule development. *Front. Microbiol.* 9, 1–18. doi:10.3389/fmicb.2018.02068.
- Weiberg, A., Wang, M., Lin, F. M., Zhao, H., Zhang, Z., Kaloshian, I., et al. (2013). Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science.* 342, 118–123. doi:10.1126/science.1239705.
- Wilson, R. C., and Doudna, J. A. (2013). Molecular mechanisms of RNA interference. *Annu. Rev. Biophys.* 42, 217–239. doi:10.1146/annurev-biophys-083012-130404.
- Xie, X., Lin, H., Peng, X., Xu, C., Sun, Z., Jiang, K., et al. (2016). Arbuscular mycorrhizal symbiosis requires a phosphate transceptor in the *Gigaspora margarita* fungal symbiont. *Mol. Plant.* doi:10.1016/j.molp.2016.08.011.
- Yang, Y., Liu, T., Shen, D., Wang, J., Ling, X., Hu, Z., et al. (2019). Tomato yellow leaf curl virus intergenic siRNAs target a host long noncoding RNA to modulate disease symptoms. *PLoS Pathog.* 15, 1–22. doi:10.1371/journal.ppat.1007534.
- Zanini, S., Šečić, E., Jelonek, L., and Kogel, K.-H. (2018). A bioinformatics pipeline for the analysis and target prediction of RNA effectors in bidirectional communication during plant-microbe interactions. *Front. Plant Sci.* 9, 1212. doi:10.3389/fpls.2018.01212.
- Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic Acids Res.* 46, D754–D761. doi:10.1093/nar/gkx1098.
- Zhang, T., Zhao, Y. L., Zhao, J. H., Wang, S., Jin, Y., Chen, Z. Q., et al. (2016). Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen. *Nat. Plants* 2. doi:10.1038/nplants.2016.153.

CHAPTER 6

6. General remarks and perspectives

6.1. General remarks and perspectives

In the recent past, some authors questioned the fact that the reductionistic approach in biology reached the limit of its own capability to explain the nature of living systems (Kaiser, 2011). This was particularly evident in molecular biology since, among all the different disciplines in biological sciences, it is the one most drenched in reductionistic and deterministic assumptions (Regenmortel, 2004; Mazzocchi, 2008). In fact, at the basis of classic molecular biology approach there is the idea that, dissecting a phenomenon into its minimal parts, it is possible to explain the phenomenon as a whole and determine the effect of the single components on the final system. However, this approach alone turned out to be incomplete since it does not take into account that biological systems, by their very nature, are both complex - so are characterized by emergent properties that can not be inferred by the study of their individual parts - and partially chaotic - so it is sometimes impossible to establish a cause-effect relationship (Regenmortel, 2004; Mazzocchi, 2008). The need to overcome this limitation has been somewhat solved with the advent of Big Data in biology (the “omics era”; Mazzocchi, 2015). However, despite their extreme capacity in describing a biological object as a whole, ‘omics’ technologies are limited in explaining the functioning of biological processes. In this context, the balance of both approaches - classic molecular biology (reductionism) and omics technologies (holism) - has given a new push to biological sciences. This was particularly evident in those topics in which the molecular tools are extremely limited, such as the case of mycorrhizal research. At present, several of the most important recent discoveries about AMS have been derived from an initial “omic observation” which led to the development of a hypothesis and finally to its validation through a classic reductionistic approach. An emblematic example was the discovery that the primary reason of AMF obligate biotrophy is the dependence on plant-derived supply of lipids (Jiang et al., 2017; Keymer et al., 2017; Luginbuehl et al., 2017). In this case, the initial observation of the lack of fatty acids synthase-related genes in AMF genomes has led to the characterization of the molecular details of a complex process.

In this thesis I tried to follow a combined reductionistic-holistic approach for the characterization of RNAi and sRNAs in AMF. The mining and analysis of already published genomics and transcriptomics data have revealed that AMF

are equipped with a very peculiar RNAi machinery, characterized by the expansion of AGO-like and RdRp families (Chapter 2 and 5). Most of these genes are expressed and some (such as AGO-like of *R. irregularis*, Chapter 2) are up-regulated in extra-radical compared to intra-radical mycelium. Why do AMF show these AGO and RdRp gene expansions? Is this a consequence of their need, in the absence of other processes, to defend from transposable element proliferation (Muszewska et al., 2017)? Are AMF AGO-like and RdRp functionally redundant or do they have specific roles? It will also be interesting to decipher the role of the genes encoding small peptides only containing the non-AGO but piwi domain or the RdRp domain which have been found (also as expressed sequence) in particular in *R. irregularis* (Chapter 2). The comparative analysis of genomes from other AMF as well as other related fungi (i.e. Mucoromycotina; Bonfante and Venice, 2020) may offer a more clear picture of the evolution and diversity of the RNAi machinery; but answering the questions posed above will be very hard if in the next future protocols for AMF stable genetic transformation will not be developed. A potential alternative or additional approach to understand the role of AMF RNAi genes could be complementation assays in *M. circinelloides*, the fungal species related to AMF for which a complete collection of RNAi mutants is available and for which the involvement of DCL, RdRp and AGO in RNAi pathway has been well characterized (Garre et al., 2014).

Small RNA-seq data produced in this work have highlighted the existence of two peculiar populations of nuclear DNA-mapping sRNA that partially differ between the two analyzed species, *R. irregularis* and *G. margarita* (Chapter 2 and 5). However, both species show a high amount of sRNA reads mapping to protein-encoding genes (more pronounced in *R. irregularis*) that has suggested, similarly to the exonic-siRNAs of *M. circinelloides* (Nicolas et al., 2010), an active role of RNAi in the regulation of the endogenous gene expression. As far as concern the number of sRNA-generating loci overlapping protein-encoding genes or overlapping intergenic regions, the difference observed between the 2 AMF could be due to the higher abundance of intergenic regions in the very large genome of *G. margarita* (Venice et al., 2019). Moreover, the fact that these sRNA populations are clearly grouped in different clusters, based on several features of the loci encoding for them, has pointed out the occurrence of different active pathways for sRNA generation in

AMF.

Another peculiar feature that has been observed in *G. margarita* is that mitochondrial DNA-derived sRNAs have structural characteristics similar to the nuclear DNA-derived -so cytosolic- ones (size of 25 nt and uracil enrichment at the 5'-end - Chapter 5). This was not observed in *Rhizophagus*. This finding raises the question about which process may be involved in the origin of these mitochondrial sRNAs and what is their function. Although the mitochondrial genome was shown to encode abundant small noncoding RNAs (Ro et al., 2013), little is known about their biological role (Vendramin et al., 2017). In mammals recent evidence indicates that non-coding RNAs (ncRNAs) may contribute to the synchronization of essential cellular and mitochondrial biological processes, acting as "messengers" between the nucleus and the mitochondria (Vendramin et al., 2017); notably, the dysregulation of these pathways may lead to aging-related diseases, including cancer. In plants the systematic analysis of mitochondrial as well as chloroplast sRNAs suggests organelle-specific mRNA stabilization mechanisms (Ruwe et al., 2016). Data from the fungal kingdom are extremely limited (Marzano et al., 2018; Shang et al., 2018) and the topic would deserves more investigations.

The results also suggest the presence of an RNAi pathway in *G. margarita* able to process sRNAs from viral sequences. Unfortunately, the nature of the relationship between the viral population and the host fungus *G. margarita* remains obscure. The *G. margarita* virome characterization also revealed the existence of DNA fragments corresponding to non-retroviral replicating RNA viruses that may play a role in an unknown host viral defense strategy (Goic et al., 2016; Chapter 4). We currently hypothesize that these DNA fragments may be originated by the activity of fungal reverse transcriptases which are supposed to be very abundant due to the presence of many retrotransposons in the *G. margarita* genome. For this purpose we are investigating the effect of azidothymidine (AZT), a reverse transcriptase inhibitor, on the viral title and the production of the corresponding viral DNA. A viruses-free strain (not easy to obtain due to the complex biology of AMF) would also be instrumental to clarify the effect of these viral infections on fungal biology, in analogy to what has been done for the endobacterium *Ca. glomeribacter gigasporarum* (Salvioli et al., 2016). Despite experimental limitations, AMF *G. margarita* currently represents an interesting model to explore the complex relationships

in a symbiotic metaorganism including a fungus, a bacterium and several viruses.

As a further results of the thesis, *in silico* analyses revealed that hundreds of AMF sRNAs - and also AMF viral-derived sRNAs - may be involved in cross-kingdom RNAi, by targeting host plant genes (Chapter 2 and 5). However, since the prediction is based on complementarity of a short sequence and considering the complexity of plant and fungal genomes, false positive results are likely to occur, even though specific software have been developed. This problem is intrinsic to results obtained from computational analyses; bioinformatics outputs are therefore used to formulate hypotheses and plan specific experiments to confirm them. Indeed, in this PhD thesis “reductionistic” wet lab analyses were used to validate some of the *in silico* predictions, and preliminary results seem promising (Chapter 3). A RNAi activity was found for a fungal sRNA towards two plant mRNAs in the artificial assays of *N. benthamiana* leaves. Other putative fungal sRNA-plant target mRNA pairs could be further analyzed in the future with the same approach. However, additional evidence must be obtained from the specific biological system, that is mycorrhizal roots. One of the most crucial experiments to prove cross-kingdom RNAi in AMS will be the characterization of sRNAs from AGO immunoprecipitated samples from mycorrhizal roots (Chapter 3). This will allow us to understand if sRNA are actually exchanged from AMF to host plants.

The work of this thesis, for the first time, has opened a window on the world of sRNA in AMF: this contribution may represent a tip of an iceberg; we still have to better understand the significance of sRNA for the AMF themselves, the potential involvement in the communication with the host plant through the so called cross-kingdom RNAi and the relationship with the viruses thriving in AMF.

6.2. References

- Bonfante, P., and Venice, F. (2020). Mucoromycota: going to the roots of plant-interacting fungi. *Fungal Biol. Rev.* doi:10.1016/j.fbr.2019.12.003.
- Garre, V., Nicolás, F. E., Torres-Martínez, S., and Ruiz-Vázquez, R. M. (2014). “Fungal RNA biology,” in *Fungal RNA Biology*, 1–395. doi:10.1007/978-3-319-05687-6.

- Goic, B., Stapleford, K. A., Frangeul, L., Doucet, A. J., Gausson, V., Blanc, H., et al. (2016). Virus-derived DNA drives mosquito vector tolerance to arboviral infection. *Nat. Commun.* 7, 12410. doi:10.1038/ncomms12410.
- Jiang, Y., Wang, W., Xie, Q., Liu, N., Liu, L., Wang, D., et al. (2017). Plants transfer lipids to sustain colonization by mutualistic mycorrhizal and parasitic fungi. *Science (80-.)*. 356, 1172–1173. doi:10.1126/science.aam9970.
- Kaiser, M. I. (2011). The Limits of Reductionism in the Life Sciences. *Hist. Philos. Life Sci.* 33, 453–476.
- Keymer, A., Pimprikar, P., Wewer, V., Huber, C., Brands, M., Bucerius, S. L., et al. (2017). Lipid transfer from plants to arbuscular mycorrhiza fungi. *Elife* 6, e29107. doi:10.7554/eLife.29107.
- Luginbuehl, L. H., Menard, G. N., Kurup, S., Van Erp, H., Radhakrishnan, G. V, Breakspear, A., et al. (2017). Fatty acids in arbuscular mycorrhizal fungi are synthesized by the host plant. *Science (80-.)*. 356, 1175–1178. doi:10.1126/science.aan0081.
- Marzano, S. Y. L., Neupane, A., and Domier, L. (2018). Transcriptional and small RNA responses of the white mold fungus *sclerotinia sclerotiorum* to infection by a virulence-attenuating hypovirus. *Viruses* 10. doi:10.3390/v10120713.
- Mazzocchi, F. (2008). Complexity in biology. Exceeding the limits of reductionism and determinism using complexity theory. *EMBO Rep.* 9, 10–14. doi:10.1038/sj.embor.7401147.
- Mazzocchi, F. (2015). Could Big Data be the end of theory in science? *EMBO Rep.* 16, 1250–1255. doi:10.15252/embr.201541001.
- Muszevska, A., Steczkiewicz, K., Stepniewska-Dziubinska, M., and Ginalski, K. (2017). Cut-and-Paste Transposons in Fungi with Diverse Lifestyles. *Genome Biol. Evol.* 9, 3463–3477. doi:10.1093/gbe/evx261.
- Nicolas, F. E., Moxon, S., de Haro, J. P., Calo, S., Grigoriev, I. V., Torres-Martínez, S., et al. (2010). Endogenous short RNAs generated by Dicer 2 and RNA-dependent RNA polymerase 1 regulate mRNAs in the basal fungus *Mucor circinelloides*. *Nucleic Acids Res.* 38, 5535–5541. doi:10.1093/nar/gkq301.
- Regenmortel, M. H. V. Van (2004). Reductionism and complexity in molecular biology. *EMBO Rep.* 5, 1016–1020. doi:10.1038/sj.embor.7400284.
- Ro, S., Ma, H. Y., Park, C., Ortogero, N., Song, R., Hennig, G. W., et al. (2013). The mitochondrial genome encodes abundant small noncoding RNAs. *Cell Res.* 23, 759–774. doi:10.1038/cr.2013.37.
- Ruwe, H., Wang, G., Gusewski, S., and Schmitz-Linneweber, C. (2016). Systematic analysis of plant mitochondrial and chloroplast small RNAs suggests organelle-specific mRNA stabilization mechanisms. *Nucleic Acids Res.* 44, 7406–7417. doi:10.1093/nar/gkw466.
- Salvioli, A., Ghignone, S., Novero, M., Navazio, L., Venice, F., Bagnaresi, P., et al. (2016). Symbiosis with an endobacterium increases the fitness of a mycorrhizal fungus, raising its bioenergetic potential. *ISME J.* 10, 130–144. doi:10.1038/ismej.2015.91.
- Shang, J., Yang, Y., Wu, L., Zou, M., and Huang, Y. (2018). The *S. pombe* mitochondrial transcriptome. *Rna* 24, 1241–1254. doi:10.1261/rna.064477.117.

Vendramin, R., Marine, J., and Leucci, E. (2017). Non-coding RNA s: the dark side of nuclear-mitochondrial communication. *EMBO J.* 36, 1123–1133. doi:10.15252/emj.201695546.

Venice, F., Ghignone, S., Salvioli, A., Amselem, J., Novero, M., Xianan, X., et al. (2019). At the nexus of three kingdoms: the genome of the mycorrhizal fungus *Gigaspora margarita* provides insights into plant , endobacterial and fungal interactions. 00. doi:10.1111/1462-2920.14827.