

Conserved Proteins of the RNA Interference System in the Arbuscular Mycorrhizal Fungus *Rhizoglyphus irregulare* Provide New Insight into the Evolutionary History of *Glomeromycota*

Soon-Jae Lee¹, Mengxuan Kong¹, Paul Harrison², and Mohamed Hijri^{1,*}

¹Département de Sciences Biologiques, Institut de Recherche en Biologie Végétale, Université de Montréal, 4101 Rue Sherbrooke Est, Montréal, QC, H1X 2B2, Canada

²Department of Biology, McGill University, Montréal, Quebec, Canada

*Corresponding author: E-mail: mohamed.hijri@umontreal.ca.

Accepted: January 8, 2018

Data deposition: This project has been deposited at GenBank under the accession MF926648 - MF926651

Abstract

Horizontal gene transfer (HGT) is an important mechanism in the evolution of many living organisms particularly in Prokaryotes where genes are frequently dispersed between taxa. Although, HGT has been reported in Eukaryotes, its accumulative effect and its frequency has been questioned. Arbuscular mycorrhizal fungi (AMF) are an early diverged fungal lineage belonging to phylum *Glomeromycota*, whose phylogenetic position is still under debate. The history of AMF and land plant symbiosis dates back to at least 460 Ma. However, *Glomeromycota* are estimated to have emerged much earlier than land plants. In this study, we surveyed genomic and transcriptomic data of the model arbuscular mycorrhizal fungus *Rhizoglyphus irregulare* (synonym *Rhizophagus irregularis*) and its relatives to search for evidence of HGT that occurred during AMF evolution. Surprisingly, we found a signature of putative HGT of class I ribonuclease III protein-coding genes that occurred from autotrophic cyanobacteria genomes to *R. irregulare*. At least one of two HGTs was conserved among AMF species with high levels of sequence similarity. Previously, an example of intimate symbiosis between AM fungus and cyanobacteria was reported in the literature. Ribonuclease III family enzymes are important in small RNA regulation in Fungi together with two additional core proteins (Argonaute/piwi and RdRP). The eukaryotic RNA interference system found in AMF was conserved and showed homology with high sequence similarity in *Mucoromycotina*, a group of fungi closely related to *Glomeromycota*. Prior to this analysis, class I ribonuclease III has not been identified in any eukaryotes. Our results indicate that a unique acquisition of class I ribonuclease III in AMF is due to a HGT event that occurred from cyanobacteria to *Glomeromycota*, at the latest before the divergence of the two *Glomeromycota* orders *Diversisporales* and *Glomerales*.

Key words: RNA interference, ribonuclease, Argonaute, Dicer, arbuscular mycorrhizal fungi, cyanobacteria, symbiosis, horizontal gene transfer, gene expression, gene evolution.

Introduction

Horizontal gene transfer (HGT) is an acquisition of genetic material between two or more organisms that may be unrelated by evolutionary descent. HGT is distinct from the vertical transmission of DNA from parent to offspring. HGTs were commonly reported in prokaryotes (Soucy et al. 2015; Pinto-Carbo et al. 2016). Although, HGTs were also reported in eukaryotic organisms including Fungi, their frequency, and

cumulative effects on eukaryotic genomes are under debate (Martin 2017; Soanes and Richards 2014; Soucy et al. 2015). The events of HGT occur more frequently, if partners live in symbiotic association (Chaib De Mares et al. 2015; Davis and Xi 2015; He et al. 2016; Pinto-Carbo et al. 2016). During evolution, partners involved in close symbiosis could have exchanged their genetic material in both directions by HGT. These transferred genes can be used to trace back their

evolutionary history (Chaib De Mares et al. 2015). Many examples of HGT between endosymbiotic prokaryotes and their hosts have been reported, including *Mollicutes-related endobacteria*, the bacterium which forms endosymbiosis with *Glomeromycota* and *Mucoromycotina* (Kuo 2015; Naito et al. 2015; Torres-Cortes et al. 2015). These studies found between 3% and 5% of protein-coding genes in *Mollicutes-related endobacterial* genomes, were horizontally transferred from the ancestral fungi of *Glomeromycota* and *Mucoromycotina*. The finding supports the hypothesis of ancestral symbiosis between a fungal ancestor and a bacterial ancestor. However, whether HGTs in *Glomeromycota* have occurred is unclear.

Mycorrhiza is a mutualistic association between fungi and plants and it is one of the most intimate and ubiquitous symbioses found on earth (Smith and Read 2008). Arbuscular mycorrhizal fungi (AMF) form mycorrhizal symbioses known as endomycorrhiza. AMF are considered as an early diverged fungal lineage assigned to phylum *Glomeromycota*, and they are obligatory symbionts that establish symbiosis with host plant root cells, where they form special structures called "arbuscules" in which nutrients are exchanged between partners (Bonfante and Genre 2010). After colonization of host cell in the cortex of roots, AMF elongate the network of hyphae along and beyond the rhizosphere. Through the hyphae, AMF absorbs various nutrients from the soil (including phosphate and nitrogen), that are used as an exchange of carbon supply on the host plant's photosynthesis (Nouri et al. 2014). Up to date, over 80% of land plant species are known to establish the symbiosis with AMF (Oehl et al. 2011). The history of AMF and autotrophic land plant's symbiosis was dated back to at least 460 Ma, according to fossil records and molecular clock analysis (Simon et al. 1993; Redecker et al. 2000). On the other hand, detailed studies by molecular clocks indicated that the appearance of *Glomeromycota* is much older than the emergence of land plant on earth (Berbee and Taylor 2010; Bidartondo et al. 2011). At the same time, it was recently reported that fungus-like mycelial fossils were found in a 2.4-billion-year-old basalt (Bengtson et al. 2017), which showed that the entire fungal lineage is much older than previously thought, and their early evolution may have occurred in deep ocean. It is still unclear at what time did the AMF ancestor arise and whether they had previous symbiotic partners before the land plants, but we may find some evidence to clarify their evolution by tracking HGT events.

The processing and expression of small noncoding RNAs (sRNAs) is one of the crucial biological processes in symbiotic organisms (Formey et al. 2016). sRNAs can originate from bioprocessing of small interference RNAs (siRNAs) or microRNAs (miRNAs) in the cell. Symbiosis involving the sRNA processing mechanism has been previously reported in other symbiotic organisms, such as in the symbiosis between soybean and *Rhizobium*, or between maize and diazotrophic bacteria (Thiebaut et al. 2014; Yan et al. 2015;

Formey et al. 2016). Branscheid et al. (2011) revealed that mycorrhizal plants actively express and regulate symbiosis-related miRNAs (microRNAs) in the establishment and development of AMF mycorrhization. However, the possible related regulation mechanism of sRNA in AMF was yet far from understanding.

In eukaryotes, the sRNA processing and regulation mechanism is called RNA interference (RNAi). RNAi and gene silencing has been reported in a broad range of eukaryotic organisms, including Fungi (Dang et al. 2011; Torres-Martinez and Ruiz-Vazquez 2016), where RNAi-related pathways play a role in various functions comprising genomic surveillance, gene regulation, and genomic defense. The study of the RNAi system in Fungi has been expanded by the discovery of a phenomenon called "Quelling" in *Neurospora crassa* (Pickford et al. 2002). Quelling is post-transcriptional gene silencing guided by sRNAs resulting from the RNAi process. The principle of the fungal RNAi process is to form sRNAs which can act as regulators for cellular processes such as development, RNA stability, and processing, host defense, chromosome segregation, transcription, and translation (Dang et al. 2011). There are three components known in the RNAi system: Dicer, Argonaute/piwi (Ago), and RNA dependent RNA polymerase (RdRP) (Nicolas and Ruiz-Vazquez 2013). Dicer and RdRP play roles in sRNA biosynthesis. RdRP produces double-stranded RNAs (dsRNAs) from single-stranded RNAs (ssRNAs) which can originate from endogenous transcripts (e.g., transposable elements) or foreign RNAs (e.g., viral RNAs). Dicer is a class IV eukaryotic ribonuclease III enzyme (BRENDA 3.1.26.3) which chops target dsRNAs into sRNAs. In Fungi, together with Dicer, an additional ribonuclease III enzyme belonging to class II which shows differences in protein architecture, also contributes to digestion of dsRNAs. The same function is mediated by class I ribonuclease III enzyme in prokaryotes (Lamontagne et al. 2000; Liang et al. 2014). sRNA biogenesis could also directly function as a defense mechanism against viral infection. Biosynthesized sRNAs are loaded to RNA-induced silencing complex (RISC) or RNA-induced transcriptional silencing (RITS) complex, which have AGOs as core components. By loading of sRNAs, RISC or RITS complex can recognize and approach to target RNAs to trigger the silencing.

Our study aims to investigate the evolutionary history of the RNAi system in *Rhizoglyphus irregularis*. We asked the question: does HGT occur between *Glomeromycota* and their associated microbes? The rationale was that HGT might have higher chance to occur between symbiotic partners particularly in genes involved in sRNA processing mechanisms during evolution. To address our question, we searched for the RNAi system in the model arbuscular mycorrhizal fungus, *R. irregularis* whose expressed sequence tags (ESTs), transcriptome and genome are publicly available, using three RNAi core genes of three model fungi *Neurospora crassa*, *Mucor circinelloides*, and *Schizosaccharomyces pombe* as queries

(supplementary table S1, Supplementary Material online). BLAST-based analyses with ESTs and transcriptomic data were conducted followed by polymerase chain reaction (PCR) and Sanger sequencing, in the case mRNA level expression was not confirmed. All of the three putative core proteins of the AMF RNAi system showed high sequence identity with those proteins of *Mucoromycotina*. Surprisingly, not only does the fungal RNAi system operate with eukaryotic (class IV) ribonuclease III proteins but we also discovered that AMF possess two additional prokaryotic (class I) ribonuclease III protein-coding genes arising by putative HGT from cyanobacteria. This HGT was not observed in *Mucoromycotina* and other basal fungi, whereas one of the two enzymes resulted from HGT was highly conserved among five different AMF species along with two different orders (*Diversisporales* and *Glomerales*) in *Glomeromycota*, showing that HGT of the enzyme-coding gene occurred at the latest in the ancestor of *Diversisporales* and *Glomerales*. The transcriptome of horizontally transferred cyanobacterial ribonuclease III enzymes were expressed in both germination and mycorrhization phases of AMF life cycle.

Materials and Methods

Sequence Identification of RNAi Core Genes

Core RNAi genes of *Neurospora crassa*, *Mucor circinelloides*, and *Schizosaccharomyces pombe*, were used as reference sequences (supplementary table S1, Supplementary Material online) to identify relevant homologs in the transcriptomic and genomic data of *R. irregularis* using TBLASTN (Altschul et al. 1997). These transcriptomic and genomic data include published expressed sequence tags (ESTs) and the genome of *R. irregularis* isolate DAOM-197198 Version 2.0 available at the JGI Genome Portal (Tisserant et al. 2012, 2013). Orthology of the *R. irregularis* genes was first assessed using the reciprocal best hits approach in conjunction with domain architecture analysis. Hits of e-value lower than $1E^{-04}$ with a query coverage $>60\%$, were compared with all proteins of *N. crassa*, *M. circinelloides*, and *S. pombe* using BLASTX. Putative open reading frames of *R. irregularis* candidate sequences were annotated by pairwise comparisons using BLASTX against GenBank sequences. Accession numbers of all sequences of *R. irregularis* used in this study, are shown in supplementary table S2, Supplementary Material online. To confirm the identity of these homologs, retrieved sequences from reciprocal BLAST were also examined for functional domains conserved in the core RNAi genes using SMART 7 (Simple Modular Architecture Research Tool) (<http://smart.embl.de/>; last accessed January 15, 2018) and InterProScan v. 61.0 (<https://www.ebi.ac.uk/interpro/>; last accessed January 15, 2018). For AGO homologs, ArgoN (IPR032474), PAZ (Piwi-Argonaute-Zwille) domain (IPR003100), and Piwi (IPR003165) domains were searched for. RNA dependent RNA polymerase domain (IPR007855) was searched for in

RDRP homologs. For DICER, ATP-binding helicase domains (DEXDc (IPR014001), HELICc (IPR001650), and ribonuclease III (RIBOc) domains (IPR000999) were investigated. Candidates were considered as hits if they contained all domains and the cut-off e-value of each domain hit was $1E^{-04}$. With retrieved candidates of *R. irregularis* as a query, the RNAi genes of a broad range of fungal taxa, including *Ascomycota*, *Basidiomycota*, and *Mucoromycotina* were searched for using EMBL databases of orthologous groups and functional annotations using EggNOG v. 4.5.1 (<http://eggnogdb.embl.de>). For the EggNOG orthologous sequence group search, we applied cut-off e-value = E^{-30} with the query coverage over 60%. Finally, to confirm homology, *R. irregularis* sequences were compared with orthologous sequences in representative Fungi and, when it is possible, with representatives of well-defined outgroups of Fungi (such as *Arabidopsis thaliana* or *Drosophila melanogaster*), using phylogeny. Full species name and related accession codes of all sequences used for phylogenetic analyses were noted with each phylogeny. Protein multiple sequence alignments (MUSCLE v. 3.8; Edgar 2004) were inspected and divergent or ambiguous positions were removed using BioEdit v. 7. For each protein, the best evolutionary models with the lowest Akaike information criterion (AIC) were determined using ProtTest 3 (Darriba et al. 2011) and the selected model for each phylogeny were noted with each phylogenetic tree. The phylogenetic trees were inferred by PhyML v. 3.0 (Guindon et al. 2010), using the best model with 1,000 bootstrap replicates.

Sequence Identification of Class I Ribonuclease III Genes

Two sequences of class I (prokaryotic) ribonuclease III genes (*rirnc 2* and *rirnc 3*) were found during the identification of ribonuclease III homologs in *R. irregularis* described in the previous section. Domain architectures were examined using SMART 7 (<http://smart.embl.de/>) and InterProScan v. 61.0 (<https://www.ebi.ac.uk/interpro/>) with each domain's cut-off e-value = $1E^{-04}$. Candidate containing both ribonuclease III (RIBOc) domain (IPR000999) and double stranded RNA-binding domain (DSRM) (IPR014720) were considered as a hit. Sequences were validated for mRNA level expression by having EST contigs as well as RNA extraction followed by PCR and Sanger sequencing described in PCR and Sanger sequencing section. The corresponding genome scaffold for the two sequences were found using BLASTN. Similar sequences of these genes were also found in the *R. irregularis* genome data published in another study using single nucleus sequencing of *R. irregularis* (coverage depth 60 \times) (Lin et al. 2014). In addition, we found other similar sequences in the GenBank database (supplementary table S2, Supplementary Material online). In order to investigate the conservation of the core catalytic (RIBOc) domain from class I ribonuclease III homologs among different species of *Glomeromycota*, a TBLASTN search was performed with protein sequences of

RIBOc in *rimc 2* and *rimc 3* as queries against the genomes of four other AMF species (*Rhizoglossum clarum* (syn, *Rhizophagus clarus*), *Rhizoglossum diaphanum* (syn, *Rhizophagus diaphanous*), *Claroideoglossum etunicatum*, and *Scutelospora calospora* of which genomic data are available in our lab. *Rhizoglossum clarum* and *R. diaphanum* were sequenced by Roche 454 and *C. etunicatum* and *S. calospora* were sequenced by Illumina Mi-Seq. Quality trimming of the raw sequences applied before the assembly that was conducted with Geneious 8.0. Contigs with coverage $<6\times$ were discarded. Hits of e-value lower than $1E^{-10}$ with the query coverage $>60\%$, were examined for domain architecture using SMART 7 (Simple Modular Architecture Research Tool) (<http://smart.embl.de/>) and InterProScan v. 61.0 (<https://www.ebi.ac.uk/interpro/>). Retrieved sequences of *Glomeromycota* were used to build protein multiple sequence alignment by (MUSCLE v. 3.8; Edgar 2004) and inspected to investigate conserved amino acid residues together with RIBOc domains of class II ribonuclease III enzymes which were known to have the most similarity with class I ribonuclease III enzymes (Liang et al. 2014). The four *Glomeromycota* sequences used in our analysis were deposited in GenBank database (accession numbers from MF926648 to MF926651).

Phylogenetic Analysis of Class I Ribonuclease III Amino Acid Sequences of *R. irregulare*

Phylogenetic analysis of core RIBOc domains of class I and II ribonuclease III enzyme was conducted. RIBOc domains from class I ribonuclease III homologs found in *R. irregulare* and its relatives as well as two well-defined class II ribonuclease III enzymes from yeast (PAC1 and RNT1), were used as queries. RIBOc domains from the homologs of each query were compared in GenBank database using strict cut-off threshold (e-value = E^{-30} and query coverage over 60%), to examine their sequence similarities. Species name and accession numbers of all sequences used in phylogenetic analyses were indicated in the phylogenetic tree. RIBOc homologs from four *Glomeromycota* species: *R. clarum*, *R. diaphanum*, *C. etunicatum*, and *S. calospora*, were also included in phylogenetic analyses. The phylogenetic trees were inferred by PhyML v. 3.0 (Guindon et al. 2010), using the best model with 1,000 bootstrap replicates.

Finally, to confirm the class I ribonuclease III sequence similarity, the full length amino acid sequences of RIRNC 2 and 3 were compared with the orthologous sequences from Bacteria using phylogeny in which a sequence of Archaea was used as an outgroup. Query sequences of RIRNC 2 and 3 and bacterial orthologs were compared with data sets from EggNOG v. 4.5.1. For the EggNOG orthologous sequence group finding, we applied cut-off e-value = E^{-30} with the query coverage over 60%. Species name and accession numbers of all sequences used to phylogenetic analyses are shown

in the phylogenetic trees, which were inferred by PhyML v. 3.0 (Guindon et al. 2010), as described earlier.

AMF Isolate and Growth Condition

The arbuscular mycorrhizal fungal isolate used in this study was *Rhizoglossum irregulare* (syn. *Rhizophagus irregularis* previously named *Glomus irregulare*), deposited in the Fungal Collection of the Department of Agriculture Canada under the accession number DAOM-197198 (Ottawa, ON). *Rhizoglossum irregulare* isolate was cultured in vitro with Ri-T-DNA transformed *Daucus carota* roots in Petri dishes containing minimal medium (M) solidified with 0.4% Gellan gum (Sigma Aldrich, Canada). The cultures were kept in an incubator at 25°C in the dark. An AMF isolate is a culture that was originally started from a single spore and subcultured by transferring 1–2 cm² of gel containing a mixture of mycorrhized roots, spores, and hyphae to a new Petri dish. Spores were harvested from four months old in vitro cultures. *Rhizoglossum irregulare* isolate DAOM-197198 spore suspension was also provided by Premier Tech Biotechnologies (Rivière-du-Loup, QC).

Culture Preparation for Germination and Mycorrhization Stages

About 60,000 fresh spores (20,000 spores per replicate) of *R. irregulare* were inoculated on the 1 mL of liquid minimal medium. The spores were germinated under dark condition in an incubator at 25°C for three weeks. Culture preparation for mycorrhization stage of the isolate of *R. irregulare* was prepared as described earlier.

Extraction of RNA and cDNA Library Construction

Spores and hyphae of *R. irregulare* were harvested in a 1.5-ml microtube and quickly frozen with liquid nitrogen. Samples were grounded using a sterilized pestle. RNA extraction and DNase treatment were performed using E.Z.N.A Fungal RNA extraction kit (Omega-Biotek, Canada) according to the manufacturer's instructions. The amount of RNA per each reaction was normalized to 100 ng for further analyses. RT-PCR was done to produce cDNA library using iScript reverse transcription kit (Bio-Rad, Canada) following manufacturer's instructions.

PCR and Sanger Sequencing

PCR was used to validate *RIRNC 2* mRNA sequences. cDNA samples produced from germinating spore were used as template. PCR primers were designed by Primer 3 software ([supplementary table S3, Supplementary Material](#) online). PCR reactions were performed using *Taq* DNA polymerase (Qiagen, Canada) following the manufacturer's recommendations. PCRs were performed in a volume of 20 µl under the following conditions: denaturation at 95°C for 2 min, followed by 35 cycles of 94°C for 30 s, 48°C for 30 s, and

70°C for 2 min. Final elongation was performed at 72°C for 10 min. PCRs were run on a Mastercycler Pro S thermocycler (Eppendorf, Canada). PCR amplicons were visualized on a 1% agarose gels stained with GelRed (Invitrogen, Canada). Successful PCR amplicons were extracted from gels, purified, and sequenced using a commercial service at the Genome Quebec Innovation Center at McGill University (Montreal, QC).

TaqMan qPCR Assay

qPCR was performed to analyze the expression level changes of fungal-origin ribonuclease III (*Ridcl 1*) and bacterial-origin ribonuclease III (*Rirnc 2*) in two different phases of *R. irregularis* life cycle (spore germination and post-mycorrhization phases). Two housekeeping genes (*18S rRNA* and *ubiquitin*) were used as reference of expression level. TaqMan probes and primer sets were designed by Primer 3 software (details are described in [supplementary table S2, Supplementary Material](#) online). Probes were labelled with FAM at the 5' end and BHQ-1 at the 3' end. Fluorescence data were collected with the ViiA 7 Real-Time PCR System (Life Technologies, Canada). qPCR was performed in three biological replicates and three technical triplicates.

Results

Conservation of RNAi System in Model AMF, *R. irregularis*

The genome of the model AMF species *R. irregularis* isolate DAOM-197198 possessed all of the three core genes (Dicer, RdRP, Argonaute/Piwi) involved in the RNAi system. RNAi core homologs retrieved using reciprocal BLAST and conserved domain searching, are summarized in [supplementary table S2, Supplementary Material](#) online, where all homologs were examined and compared using BLAST searches against the GenBank database. All homologs of RNAi system core proteins in *R. irregularis* showed high sequence similarity with orthologs of Mucoromycotina fungi. [Table 1](#) summarizes the number of homologs for each core protein of *R. irregularis* and three query species (*N. crassa*, *M. circinelloides*, and *S. pombe*) with the comparison of their genome sizes. *Rhizoglyphus irregularis* possessed the largest genome of 136.81 Mb which is 3.3-fold and 9.7-fold larger than the genomes of *N. crassa* (41.04 Mb) and *S. pombe* (14.1 Mb), respectively. *Rhizoglyphus irregularis* had a large number of homologs of Argonaute proteins (26 homologs). On the contrary, despite its large genome size, *R. irregularis* did not show any difference in number of homologs of RdRP (3 homologs) compared with phylogenetically close queries (*N. crassa* and *M. circinelloides*). Interestingly, all of the three query species had two eukaryotic ribonuclease III homologs (two IV ribonuclease III homologs in the case of *N. crassa* and *M. circinelloides* and one class II ribonuclease III with one class IV ribonuclease III in the case of *S. pombe*), whereas only one

eukaryotic ribonuclease III homologue (named as RIDCL 1, from *R. irregularis* dicer-like protein) was detected in *R. irregularis* genome.

We analyzed the domain architecture of each protein homolog. [Figure 1a–c](#) shows the conserved domains in representative homologs of each protein of the fungal RNAi system in *R. irregularis*. All of the homologs of the fungal RNAi system in *R. irregularis* ([fig. 1a–c](#)) showed protein domain architectures identical to those of the original query sequences of *N. crassa*, *M. circinelloides*, and *S. pombe*.

To investigate the detailed protein homology, we inferred phylogenies for the three core genes using Maximum Likelihood. [Figure 2](#) and [supplementary figure S1, Supplementary Material](#) online, show homology of three core proteins of RNAi system in *R. irregularis* with other fungal species including Ascomycota, Basidiomycota, and Mucoromycotina. As expected, all of the three putative core proteins of RNAi system found in *R. irregularis* clustered together with Mucoromycotina in a single clade in each phylogenetic tree. Sequences of *R. irregularis* showed high sequence similarity with orthologs of the mucoromycete *M. circinelloides* (bootstrap values: 624/1,000 for AGO, 966/1,000 for RDRP, and 1,000/1,000 for RIDCL 1). On the contrary, the orthologs from *N. crassa* and *S. pombe* showed relatively low sequence similarity with *R. irregularis* and *M. circinelloides*, therefore they were distributed in different clades ([fig. 2](#)). Two sub-clusters were observed in the cluster of AGO homologs from *M. circinelloides* and *R. irregularis* (624/1,000) ([fig. 2a](#)). Twenty-two sequences of *R. irregularis* clustered within three orthologs of *M. circinelloides* (1,000/1,000). However, four homologs of *R. irregularis* clustered separately from the 22 AGO homolog sequence. In RDRP ([fig. 2b](#)), all of three homologs from *R. irregularis* clustered with an orthologue from *M. circinelloides* (966/1,000). However, two other orthologs of *M. circinelloides* clustered separately. In DICER homologs ([fig. 2c](#)), two homologs of *M. circinelloides* and one homolog of *R. irregularis* clustered in a single clade (1,000/1,000), showing their high sequence similarity. Apparently, with a divergence of the number of homologs and amino acid residues, all of the AMF RNAi core gene candidates showed the highest similarities with proteins of Mucoromycotina, a basal fungal subphylum suggested to be the closest neighbor of Glomeromycota in other studies using mitochondrial genomes (Nadimi et al. 2012) and SSU rDNA sequences (Redecker et al. 2013). Our findings are in line with GenBank BLAST best-hit results of *R. irregularis* RNAi core protein sequences ([supplementary table S2, Supplementary Material](#) online).

Existence of Two Class I Ribonuclease III Enzymes in *R. irregularis*

Surprisingly, we found two additional homologs of ribonuclease III (RIRNC 2 and RIRNC 3) in *R. irregularis*, using reciprocal

Table 1

The Number of Homologs of Core Genes of RNAi System in *Rhizoglyphus irregularis* and Three Query Species (*Neurospora crassa*, *Schizosaccharomyces pombe*, and *Mucor circinelloides*) with Their Genome Sizes between Brackets in Mega Base Pairs (Mb)

Name of Core Component of RNAi System	The Number of Homologs in Species (with genome size published)			
	<i>R. irregularis</i> (136.81 Mb)	<i>N. crassa</i> (41.04 Mb)	<i>S. pombe</i> (14.1 Mb)	<i>M. circinelloides</i> (36.6 Mb)
RNA dependent RNA polymerase (RDRP)	3	3	1	3
Argonaute/Piwi Protein (AGO)	26	2	1	3
Ribonuclease III	1 (Class IV) 2 (Class I)	2 (Class IV)	1 (Class IV) 1 (Class II)	2 (Class IV)

NOTE.—The number of copies in *R. irregularis* was determined based on BLAST and protein structure analyses. *Rhizoglyphus irregularis* genome showed an expansion of Argonaute/Piwi genes (26 copies), and it had two class I ribonuclease III (prokaryotic) enzymes and only one class IV ribonuclease III (eukaryotic) enzyme.

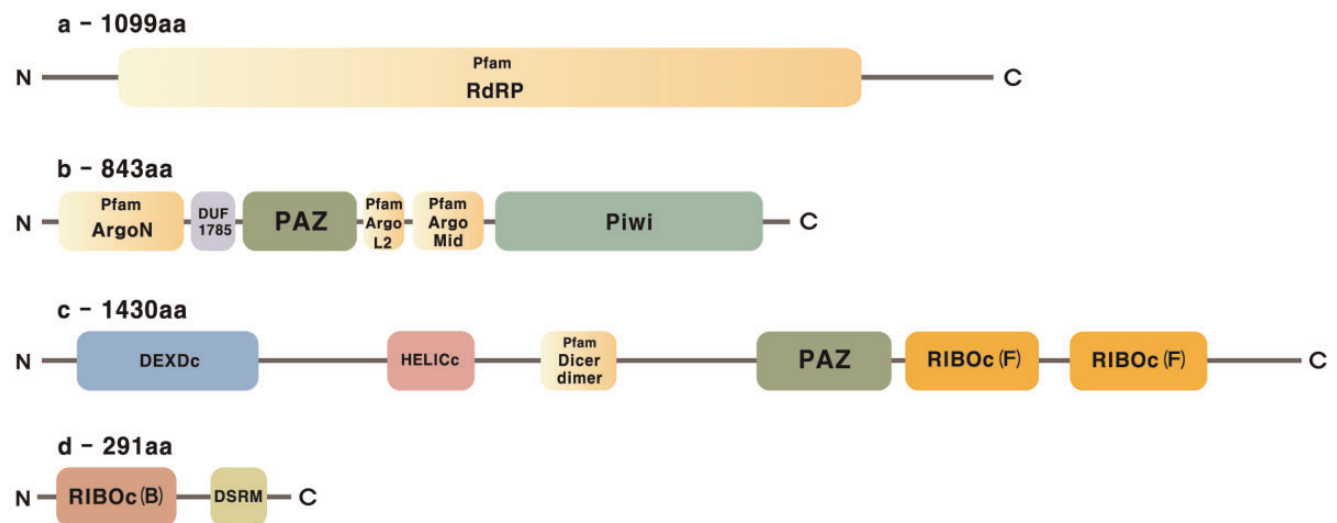


FIG. 1.—Domain based protein structures of three core genes forming RNAi system in *Rhizoglyphus irregularis*. Four representative sequences of three core proteins (RdRP [a], Argonaute [b], and Dicer [c]) were analyzed by SMART (Letunic et al. 2015) and InterProScan v. 61.0. (a) RNA dependent RNA polymerase domain (IPR007855) was found in *R. irregularis* RdRP. (b) Putative AGO homologs, N terminal domain of argonaute protein (ArgoN, Pfam domain) was detected with PAZ (Piwi-Argonaute-Zwille) domain (IPR003100) in the middle and Piwi domain (IPR003165) C terminal. (c) Class IV ribonuclease III enzyme (RIDCL) contains at its N terminal an ATP-binding helicase domains (DEXDc [IPR014001] and HELICc [IPR001650]), PAZ (IPR003100), and two catalytic ribonuclease III (RIBOc) domains (IPR000999) in C terminal. (d) Class I ribonuclease III enzyme (RIRNC) contains one ribonuclease III (RIBOc) domain (IPR000999) starting at 18 a.a. from N terminal and one double stranded RNA-binding domain (DSRM) (IPR014720) in C terminal.

BLAST (supplementary table S2, Supplementary Material online). RIRNC 2 and 3 showed relatively low e-values and low sequence similarities when they were compared with their reciprocal hits as queries (supplementary table S2, Supplementary Material online). Moreover, both of best BLAST hits of RIRNC 2 and RIRNC 3 were cyanobacterial ribonuclease III, whereas RIDCL1 had mucoromycete ortholog as the best BLAST hit. We analyzed the domain architecture of each protein homolog and we found that RIRNC 2 and RIRNC 3 have the typical protein architecture (fig. 1d) of prokaryotic ribonuclease III reported in bacteria (Class I ribonuclease III) (Court et al. 2013; Liang et al. 2014). It has been known that Fungi only possess two variants of ribonuclease III (class II and IV), which are clearly distinguished by the protein architecture from class I ribonuclease III in prokaryotes (Lamontagne et al. 2001; Court et al. 2013; Liang et al. 2014).

Interestingly, the structurally closest variant of eukaryotic ribonuclease III (PAC1) was detected in query species, *S. pombe* (supplementary table S2, Supplementary Material online). In a previous study, PAC1 showed the highest structural homology with class I ribonuclease III among other eukaryotic ribonuclease III by having one ribonuclease III (RIBOc) domain and one double stranded RNA-binding domain (DSRM) C-terminal to it (Lamontagne et al. 2000). However, PAC1 and related homologs are clearly different in domain architecture from a typical class I ribonuclease III (RIRNC 2 and 3) detected in *R. irregularis* and other prokaryotes (fig. 1d) by having an additional domain in the N-terminus (NTD domain, PDB ID: 400g), which was proven to be crucial for the enzyme's function (Lamontagne et al. 2001). PAC1 is classified as class II eukaryotic ribonuclease III with similar proteins found in other yeasts (RNT1) (Liang et al. 2014).

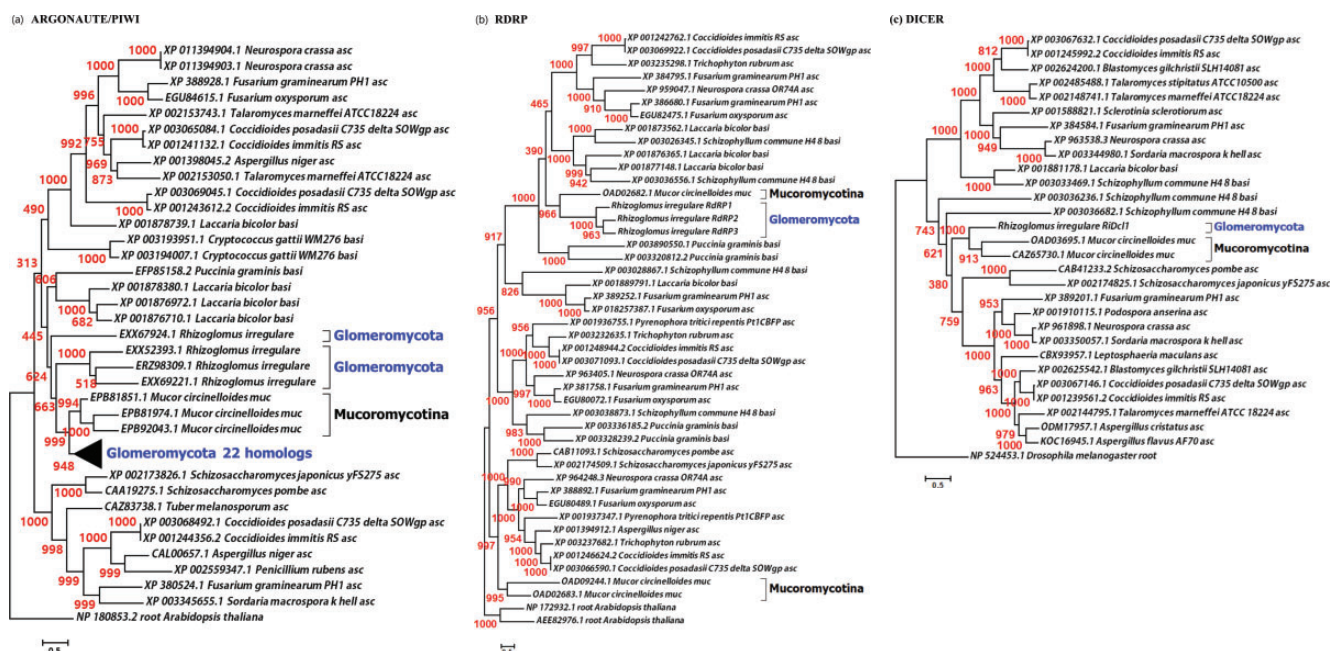


FIG. 2.—Phylogeny of three core proteins of RNAi system in AMF with other fungal species. Maximum likelihood of amino acid sequences of AGO (a), RDRP (b), and DICER (Class IV ribonuclease III) (c) were analyzed with the WAG +I+G (with four distinct gamma categories) phylogenetic model which showed the lowest AIC value. The *Rhizoglossus irregulare* (synonym of *Rhizophags irregularis*) and *Mucoromycotina* sequences were annotated in right side of taxon names, respectively. The numbers at branches correspond to bootstrap support values generated with 1,000 bootstrap replicates. All of the putative *R. irregulare* AGOs, RDRPs, and DICER showed the most sequence similarities with orthologs of *Mucor circinelloides* (624/1,000 for AGO, 966/1,000 for RDRP, and 1,000/1,000 for RIDCL 1) by forming single clade in each analysis. Trees (a) and (b) were rooted using *Arabidopsis thaliana*, whereas tree (c) was rooted using *Drosophila melanogaster*.

In *R. irregulare* genome, we found in total three ribonuclease III homologs which were classified into two types: class I (RIRNC 2 and 3) and class IV (RIDCL I). Full protein architecture between eukaryotic (class IV ribonuclease III) and prokaryotic (class I ribonuclease III) homologs found in *R. irregulare* were apparently different (fig. 1c and d). However, to clarify the differences between core RIBOC domains of class I (RIRNC2 and 3) and class IV (RIDCL1) ribonuclease III enzymes found in *R. irregulare*, we further investigated their RIBOC domains. The amino acid sequences of RIBOC domains of RIRNC 2 and 3 and RIDCL 1 of *R. irregulare* were used as queries in BLASTP searches that were performed with the entire non-redundant (nr) protein database of GenBank to investigate the taxonomical distributions of their hits. As expected, we found clear differences between class I and class IV RIBOC groups of *R. irregulare* (supplementary table S4 and fig. S2, Supplementary Material online). The taxonomical results of BLASTP of two class IV RIBOC domains (domains A and B, fig. 1c) from RIDCL 1 were 100% eukaryotic sequences (domain A: 1,304/1,304 hits and domain B: 1,368/1,368 hits). In contrast, 100% of the BLAST hit results of both class I RIBOC domains (RIRNC 2 and 3) were prokaryotic sequences (RIRNC 2: 9,779/9,779 hits, RIRNC 3: 9,515/9,515 hits). This is an evidence showing that class I and IV RIBOC domains from *R. irregulare* ribonuclease III homologs are clearly different in amino acid

sequence level, which is in agreement with previous reports (Lamontagne et al. 2001; Court et al. 2013; Liang et al. 2014). To identify the full length protein homologs of class I ribonuclease III (RIRNC 2 and 3), BLASTP-based analysis was conducted with translated *rirc2* and *rirc3* sequences (RIRNC 2 and 3, later, identical proteins were also found in *R. irregulare* proteome data set) as queries (supplementary table S5 and fig. S3, Supplementary Material online). The query coverage of top 100 hits ranged from 83% to 98% (RIRNC 2) and 55% to 87% (RIRNC 3) and e-values of top 100 hits were lower than $5E^{-38}$ (RIRNC 2) and $3E^{-35}$ (RIRNC 3). Expectedly, 100% of top 5,000 hits of BLASTP results for both RIRNC 2 and 3 were class I prokaryotic ribonuclease III enzymes, and top 100 hits of BLASTP results were mostly from cyanobacteria clades, implying the high level of sequence similarity between homolog from *R. irregulare* and cyanobacteria. To diagnose the possibilities of mis-identification of RIRNC 2 and 3 homologs, we did detailed inspection of genes of RIRNC 2 and 3 by cross-checking all of the sequence data sets publicly available for *R. irregulare* whose genome was sequenced two times in different studies using different approaches (Tisserant et al. 2013; Lin et al. 2014). Later, Tisserant et al. (2013) resequenced and updated the data set of *R. irregulare* genome (version 2.0 in JGI) with a coverage of 50x. The RIRNC 2 and 3 sequences we found in two versions 1.0 and 2.0 data sets published by

(Tisserant et al. 2013) as well as in the data set reported by (Lin et al. 2014) where the genome sequencing was done using four single nuclei of *R. irregularis* (coverage of 60×). Moreover, we retrieved hits of ESTs of *R. irregularis* published in previous studies (Tisserant et al. 2012, 2013). Both coding genes of RIRNC 2 and 3 were located in assembled genome scaffolds of *R. irregularis* genome version 2.0 (Tisserant et al. 2013). The coding gene of RIRNC 2 was located from 722,662 to 723,453 bp in genome scaffold 4 (scaffold length: 939,913 bp in version 2.0 of *R. irregularis* genome available at JGI website) without intron, whereas *rimc 3* was located from 391,302 to 392,174 bp of genome scaffold 82 (length: 423,271 bp, in version 2.0) without intron. The well-defined neighboring genes in both sides (5' upstream and 3' downstream) of *rimc 2* and *3* were also investigated, respectively. Phylogenetic analysis was conducted together with BLAST searches in GenBank of each sequence (threshold e-value = E^{-20} and query coverage over 80%). The phylogenetic analysis of each gene with comparison of other prokaryotic or eukaryotic orthologs as outgroups showed that all of the tested genes were typical fungal genes (supplementary fig. S4, Supplementary Material online). Taken together, these evidence clearly confirmed that class I homolog sequences are inserted in *R. irregularis*'s genome.

Expression Confirmation of Class I Ribonuclease III Enzyme

We found expression evidence of *rimc 3* in EST data sets. However, at the time with initial EST data set of *R. irregularis*, we did not find any transcripts of *rimc 2* matching with these EST libraries. We therefore experimentally verified its expression status using RT-PCR followed by Sanger sequencing on *R. irregularis* mRNA. The primer set was designed to target open reading frame (ORF) region of *rimc 2* sequence (supplementary table S1, Supplementary Material online). We amplified a partial sequence of *rimc 2* (110 bp) from cDNA of *R. irregularis*. The obtained sequence was identically matched with candidate sequence of *rimc 2* retrieved by sequence analysis. The updated *R. irregularis* EST data set available at the JGI had an identical match with *rimc 2*, confirming the expression of the gene. Furthermore, we tested the gene expression of two types of ribonuclease III family enzymes (class IV ribonuclease III [*ridcl 1*] and class I ribonuclease III [*rimc 2*]) in *R. irregularis* under two different life cycle stages (spore germination stage and mycorrhization stage) of AMF. Relative expression of RNA was tested with qRT-PCR using TaqMan assay and two housekeeping genes (*18s rna* and *ubiquitin*). The primers and probes were designed and successfully tested (supplementary table S2, Supplementary Material online). Figure 3 shows the different levels of the expressions of *ridcl 1* and *rimc 2* in *R. irregularis* isolate DAOM197198. The patterns of normalized results using both housekeeping genes as reference were similar. Interestingly, the expression of the gene encoding class I type ribonuclease III enzyme (*rimc 2*) was significantly higher

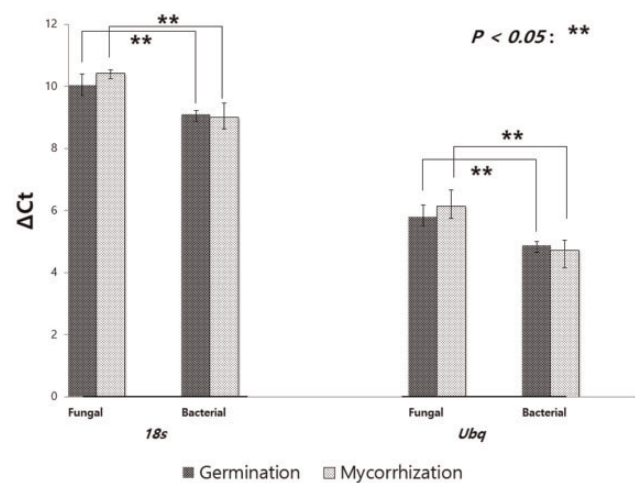


Fig. 3.—Expressions of class I (bacterial) and class IV (fungal) ribonuclease III family proteins coding mRNA (*ridcl 1* and *rimc 2*) in two different phases of *Rhizoglyphus irregularis* life cycle. Relative quantification of gene expression of *ridcl 1* and *rimc 2* using TaqMan assay and two different housekeeping genes (*18s rna* (left), *ubiquitin* (right)), in germination and mycorrhization phases. Y-axis shows threshold cycle difference between dicer genes and reference genes (calculated by $C_{tTarget} - C_{tReference}$ and converted to fold changes by formula $2^{-\Delta Ct}$). Black bars represent spore germination phase, whereas gray bars represent mycorrhization phase. The error bar indicates minimum and maximum values among the replicates. The statistical significance ($P < 0.05$) was marked on figure with star.

than the expression of the gene encoding class IV type ribonuclease III enzyme (*ridcl 1*) during both spore germination (~1.9-fold increase) ($P = 0.024253$ (with reference gene: *18s rna*) and $P = 0.027964$ (with *ubiquitin*), Student's two-tailed *t*-test) and mycorrhization stages (~2.7-fold increase) ($P = 0.017255$ (*18s rna*) and $P = 0.021482$ (*ubiquitin*), Student's two-tailed *t*-test). The difference between expressions of *rimc 2* and *ridcl 1* becomes larger in mycorrhization phase compared with spore germination phase. However, the expression level differences of same gene under different life cycle stages were not significant in our experiments.

Unique and High Level of Conservation of Cyanobacterial Ribonuclease III Enzymes in AMF

Up to date, there has been no report of the existence of class I ribonuclease III in Eukaryotes (Lamontagne et al. 2001; Liang et al. 2014). To confirm our result, we performed repeated sequence searches in all of the fungal genome and proteome databases available including GenBank and we did not find any evidence of the existence of the class I ribonuclease III enzyme coding genes (*rimc 2* and *3*) in any fungal genome except that of *R. irregularis*. No homologous sequence was found even in the basal fungal subphylum *Mucoromycotina*, which yielded the closest homology for all three fungal core proteins related with RNAi. Because class I type ribonuclease III

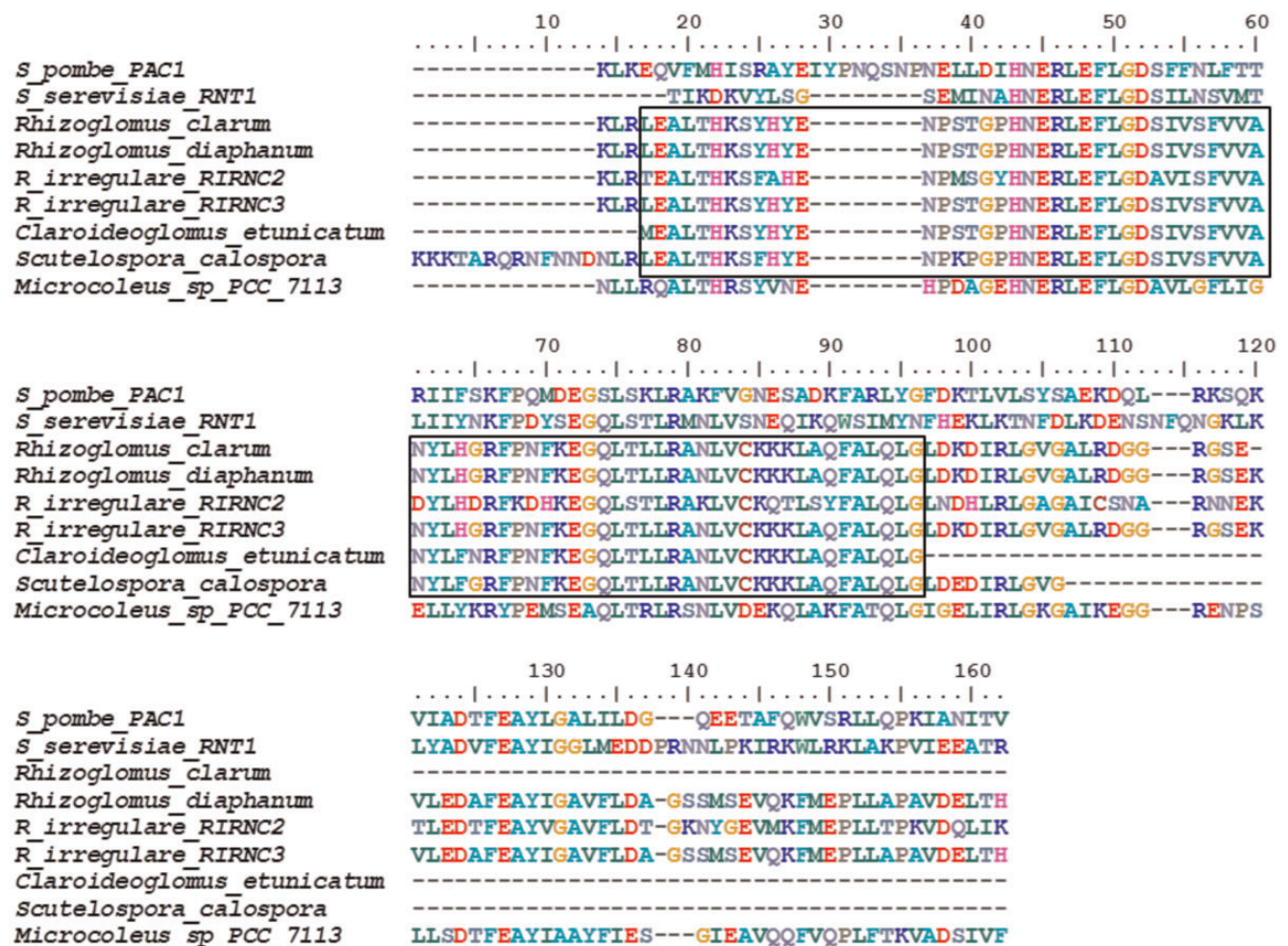


Fig. 4.—Multiple Sequence Alignment of conserved RIBOc domain from class I ribonuclease III among Glomeromycota and cyanobacteria, comparing with those of the protein structurally closest eukaryotic class II ribonuclease III from yeasts. Protein sequences from five Glomeromycota species (*Rhizoglossum irregulare*, *Rhizoglossum clarum*, *Rhizoglossum diaphanum*, *Claroideoglossum etunicatum*, *Scutelospora calospora*), two yeast species (*Schizosaccharomyces pombe* and *S. cerevisiae*) and top hit of BLASTP analysis (*Microcoleus* sp. PCC7113 (cyanobacteria), $4E^{-38}$) were used. Similar amino acid residues were marked with same color. Glomeromycota species were labeled with black rectangle (from residue 17 to 96, 80 a.a.). Amino acid residue Glu-Phe-Leu-Gly-Asp (from 48 to 52 a.a. in the figure) were highly conserved among all the species analyzed so far for class I ribonuclease III (total 52 species).

enzyme encoding genes were uniquely found in *R. irregulare*, we investigated whether catalytic domains (RIBOc, ribonuclease III domain) of RIRNC are in other AMF species (fig. 4). Four other AMF species (*Rhizoglossum clarum* (syn, *Rhizophagus clarus*), *Rhizoglossum diaphanum* (syn, *Rhizophagus diaphanous*), *Claroideoglossum etunicatum*, *Scutelospora calospora*) of which genomic data are available in our lab (data not published) were analyzed. We were unable to detect the hits of RIBOc domain of RIRNC 2 because of the incompleteness of other species' genome data sets. On the other hand, partial sequences of RIBOc domain of RIRNC 3 were detected in each of the four AMF species genome. All of four sequences of *rirc 3* homologs were deposited and are publicly available in GenBank under the accession numbers MF926648-MF926651. Figure 4 shows conserved RIBOc domain of class I ribonuclease III enzymes from different AMF species

(> 80 a.a.) with comparison of same domains in class II ribonuclease III enzymes of *S. pombe* and *S. cerevisiae*, which were reported to be the most similar fungal ribonuclease III enzymes to class I ribonuclease III enzymes (Liang et al. 2014). The shared top hit from BLASTP analysis of *R. irregulare* class I ribonuclease III enzyme (*Microcoleus* sp. PCC7113 [cyanobacteria], E-value = $2E^{-34}$ [in BLAST result of RIRNC 2] and $4E^{-38}$ [in BLAST result of RIRNC3]), was also compared for their sequence similarity). Interestingly, amino acid residues Glu-Phe-Leu-Gly-Asp (from 48 to 52 a.a. in fig. 4) were highly conserved among all of the species analyzed so far for bacterial ribonuclease III (total 52 species). AMF species sequences showed high levels of sequence similarities between each species compared with their similarity with yeasts or cyanobacteria. Three species of the *Rhizoglossum* genus showed 100% identical matching along the amino acid residues

from 17 to 96, whereas *C. etunicatum* (at a single point, two amino acid residues) and *S. calospora* (at three points, four amino acid residues) showed minor divergences. *rirc* 2 of *R. irregulare* showed a higher total number of differences at amino-acid positions (22/80) compared with *rirc* 3 of *R. irregulare*, whereas highest total differences at amino-acid positions for a *rirc* 3 homolog was 4/80 for *rirc* 3 from *S. calospora*. We further conducted detailed phylogenetic analysis with these RIBOc domains from class I and class II ribonuclease III (supplementary fig. S5, Supplementary Material online). All of the class I ribonuclease III sequences were clearly separated from those of class II ribonuclease III enzyme (high bootstrap supporting: 994/1,000). The sequence variation was also observed in the core catalytic domain (RIBOc). Taxonomical BLAST results described earlier (supplementary table S5 and fig. S3, Supplementary Material online), along with phylogenetic and amino acid sequence analyses support that discovery of RIRNC 2 and 3 in *R. irregulare* was not a result of false detection of class II ribonuclease III homologs. In supplementary figure S5, Supplementary Material online, class I ribonuclease III sequences of Glomeromycota taxa were clustered as a monophyletic clade with high level of sequence similarity and they separated from other class I ribonuclease III sequences (927/1,000). Inside of the monophyletic clade of Glomeromycota sequences, internal monophyletic clustering of RIRNC 3 homologs was observed, apart from RIRNC 2 sequence from *R. irregulare* (908/1,000). Three species of the *Rhizoglossum* genus (*R. irregulare*, *R. clarum*, *R. diaphanum*) showed high percentage of sequence similarity of RIBOc domain of RIRNC 3, and they formed a phylogenetic clade (738/1,000), distinguished from sequences of *C. etunicatum* and *Scutelospora calospora* which belong to different families in Glomeromycota.

We further inferred a phylogenetic tree of full length sequence of RIRNC 2 and 3 of *R. irregulare* and those of class I ribonuclease III orthologs from Bacteria, where Archaea sequence was used as an outgroup (fig. 5). RIRNC 3 homologous sequences from other AMF species were excluded from this analysis because they were partial sequences (supplementary fig. S5, Supplementary Material online). As expected, both RIRNC 2 and 3 clustered with Cyanobacteria class I ribonuclease III sequences with a high bootstrap support (945/1,000). Within the *R. irregulare*-cyanobacterial group, RIRNC 2 and 3 formed a single clade which showed 100% bootstrap support (fig. 5). This result supports that the event of HGT was ancient.

Discussion

Conservation of RNAi System in the Model AMF Genome

The RNAi system was highly conserved in the model AMF *R. irregulare* with respect to all three of its core proteins (Dicer, RdRP, and Agos). Three core proteins showed high levels of

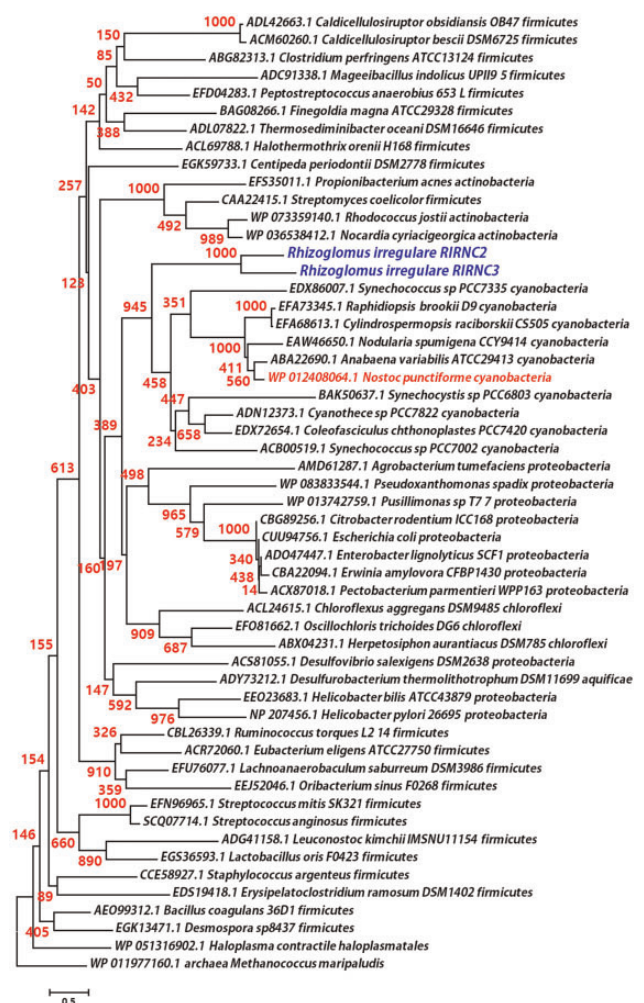


Fig. 5.—Phylogenetic analysis of class I ribonuclease III enzymes in of *Rhizoglossum irregulare*. Maximum likelihood of amino acid sequences of class I ribonuclease III enzymes were analyzed with the LG + I+G (with four distinct gamma categories) phylogenetic model which showed the lowest AIC value. The tree was rooted with class I ribonuclease sequence from archaea (*Methanococcus maripaludis*). The *R. irregulare* sequences were colored in blue. A cyanobacterial species (*Nostoc punctiforme*) which were known to form symbiosis with one basal AMF species (*Geosiphon pyriforme*) was colored in red. Cyanobacteria sequences were annotated in right side of taxon names. The numbers at branches correspond to bootstrap support values generated with 1,000 bootstrap replicates. Both RIRNC 2 and 3 of *R. irregulare* strongly formed single clade with orthologs of cyanobacteria group with high bootstrap supporting (945/1,000), clearly separated from protein groups of other bacteria.

sequence similarity with orthologs from *Mucoromycotina*, a basal fungal subphylum which was suspected to be a closely related with *Glomeromycota* in other studies (Nadimi et al. 2012, 2016; Redecker et al. 2013).

It has been documented that natural selection ensures that each organism is pressured to optimize its RNAi system to counter the particular challenges it faces (Ketting 2011). In spite of the close homology of core proteins with sister

subphylum Mucoromycotina, RNAi system of *R. irregularis* had two interesting points. First, as shown in table 1, the *R. irregularis* genome harbors large numbers of AGO homologs compared with the general fungal species including the model fungus *N. crassa* (24 more in number), or its closest neighbor mucoromycete *M. circinelloides* (23 more in number) (Dang et al. 2011; Torres-Martinez and Ruiz-Vazquez 2016). In contrast to the cases in noneukaryotic genomes, eukaryotic genome size and the number of genes are not correlated (Hou and Lin 2009). Thus, the expansion of the number of AGO homologs could not be simply explained by the large genome size of *R. irregularis*. Different with the case of AGO, the number of homologs of the other crucial core protein (RDRP) did not show differences compared with *N. crassa* or *M. circinelloides*. The number of AGO homologs in the *R. irregularis* genome was 26, close to that reported in the model nematode worm, *Caenorhabditis elegans*, which had 27 homologs of AGOs (Yigit et al. 2006). Interestingly, in the case of *C. elegans* and its relative nematodes, AGO diversity was related with symbiosis since there was additional AGO protein diversity found in parasitic nematodes compared with free-living relatives (Buck and Blaxter 2013). The variety of symbioses, including parasitism, require complex sRNA regulation in both symbiotic partners, and AGO is one of the crucial enzymes in this process (Katiyar-Agarwal and Jin 2010; Weiberg et al. 2015; Formey et al. 2016). Considering its complex symbiotic relationship with other fungi, bacteria, and obligatory symbiosis with a plant host, we hypothesized that the high number (26) of AGO homologs in *R. irregularis* reflects the complex sRNAs regulation mechanism related with symbiosis, as in nematode worms. Further investigation of protein function of each AGO of *R. irregularis* related with its symbiosis will shed light on this hypothesis. At the same time, it has been shown that *R. irregularis* isolates harbor abnormally high number of gene copies compared with other fungi, at least 76 homologs of Mating-type-high-mobility group (*Sex M* and *Sex P* genes) (Halary et al. 2013; Riley et al. 2014). These copies also showed a high intraspecific genetic polymorphism among *R. irregularis* isolates. Riley et al. (2014) explained that intragenomic gene duplications play an important role in the expansion of copy number of Mating-type-high-mobility group in *R. irregularis* isolates. Thus, we hypothesize that the high copy number of AGO genes in *R. irregularis* could also be explained by gene duplications originating from intrainolate genetic polymorphism in Glomeromycota (Hijri and Sanders 2005; Stukenbrock and Rosendahl 2005; Marleau et al. 2011; Riley et al. 2014).

The second interesting feature observed in *R. irregularis* RNAi system is the number of DICER (class IV ribonuclease III) homologs. In contrast to the large number of AGO homologs, *R. irregularis* genome only harbored a single class IV ribonuclease III homolog (RIDCL 1), similar to the cases of *S. cerevisiae* or *S. pombe* genomes. However, in yeasts, an

additional variant of eukaryotic class II ribonuclease III that digests double strand RNA was reported (Lamontagne et al. 2001). We could not find any class II homologs in *R. irregularis*. Strikingly, instead, we discovered two additional class I ribonuclease III homologs (RIRNC 2 and 3) in the *R. irregularis* genome. Domain architectures of RIRNC 2 and 3 were identical with typical prokaryotic class I ribonuclease III enzymes which have the function to digest dsRNA and they are involved in viral defense mechanism as well as transcriptome processing in Bacteria (Nicholson 1999). ESTs based sequence analysis and Sanger sequencing confirmed the expression of all class I and class IV ribonuclease III mRNAs (*ridcl 1*, *rirc 2*, and *rirc 3*) in *R. irregularis*. It is well known that class I ribonuclease III has a function in dimeric form, whereas class IV ribonuclease III acts as monomeric form (Lamontagne et al. 2001; Court et al. 2013). RT-qPCR experiment showed that a representative of class I ribonuclease III mRNA (*rirc 2*) had higher expression level than the class IV ribonuclease III mRNA (*ridcl 1*) in two contrasting phases of *R. irregularis* life cycle (spore germination and mycorrhization). The detected expression level difference could be due to the homodimer characteristic of class I ribonuclease III which needs double of the population of protein chains for forming homodimer to have catalytic function compared with monomeric class IV ribonuclease III (Sun and Nicholson 2001). At the same time, the heterodimerization of ribonuclease III in *E. coli* was also reported and demonstrated (Conrad et al. 2002; Meng and Nicholson 2008). The further investigation of protein level interaction between two class I ribonuclease III (RIRNC 2 and 3) in *R. irregularis* is required for understanding their function. The expression levels difference between class I and class IV ribonuclease III mRNAs increased in symbiosis phase (fig. 3).

Considering the core gene conservation and related two interesting features described earlier, we propose a putative mechanism of RNAi system in the model AMF, *R. irregularis* (fig. 6). Based on known functions in prokaryotic orthologous (Nicholson 1999; Lamontagne et al. 2001; Knip et al. 2014), we hypothesize that class I ribonuclease III enzymes of *R. irregularis* are also functional in the processing of dsRNA originated from exogenous RNAs or endogenous transcripts. sRNAs can be generated by class I and class IV ribonuclease III enzymes (RIDCL 1, RIDCL 2, or RIDCL 3) with or without putative RDRPs. This sRNA biogenesis could directly function as a defense mechanism against viral attack by degrading their dsRNA. Biosynthesized sRNAs could be loaded to the RISC or RITS complexes, which have putative AGOs as core components. By loading of sRNAs, RISC or RITS complex can recognize and approach to target RNAs which are complementary to guide sRNAs to trigger the silencing.

Class I Ribonuclease III Found in *R. irregularis* and HGT Event

So far, there has been no report for the existence of class I (prokaryotic) ribonuclease III in Eukaryotes (Lamontagne et al.

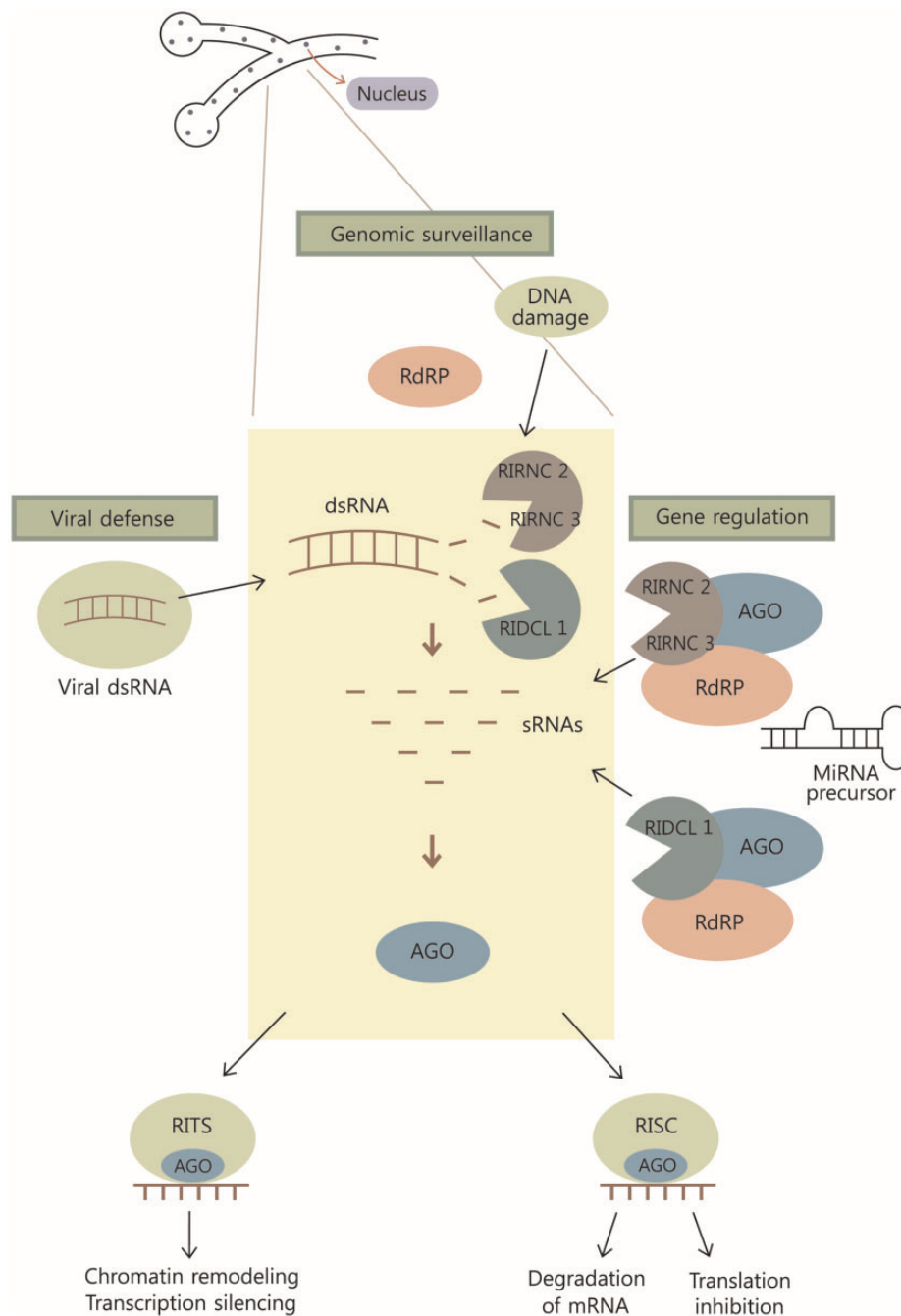


Fig. 6.—Schematic diagram of *Rhizoglyphus irregularis* putative RNAi system consisted on three core proteins. This diagram was built based on the study of RNAi system in the ascomycete *Neurospora crassa* (Dang et al. 2011). Two different types of ribonuclease III proteins were noted as RIRNC 2 and 3 (class I ribonuclease III) and RIDCL 1 (class IV ribonuclease III).

2001; Liang et al. 2014). In the kingdom Fungi, the closest type of class I ribonuclease III enzyme is class II which was reported in *S. cerevisiae* (RNT1) and *S. pombe* (PAC1) (Lamontagne et al. 2000). RNT1 and PAC1 showed structural homology with class I ribonuclease III by having one ribonuclease III (RIBOc) domain and one double stranded RNA-binding domain (DSRM) in C-terminal. However, RNT1 and

PAC1 are different in domain architecture of class I ribonuclease III by having an additional domain in the N-terminus (NTD, PDB ID: 400 g), which was proven to be crucial for the enzyme's function (Lamontagne et al. 2001). RIRNC 2 and 3 found in *R. irregularis* were clearly distinguished from RNT1 or PAC1 by lacking eukaryotic N-terminal domain and they showed identical protein architecture of prokaryotic class

I ribonuclease III. Moreover, the RIBOC domains of class II and class I ribonuclease III were also clearly distinguished (supplementary fig. S5, Supplementary Material online). Thus, we concluded that RIRNC 2 and 3 encoding genes found in *R. irregulare* are likely acquired by HGT from a prokaryotic ancestor related to Cyanobacteria (supplementary table S2, Supplementary Material online and fig. 5). At the same time, even though both RIRNC 2 and 3 showed their high level of sequence similarity with cyanobacterial ribonuclease III enzymes, respectively, they were different in length (RIRNC 2—264 a.a., RIRNC 3—291 a.a.), and sequence identity is only 52% between each. Moreover, the core catalytic domains (RIBOC) of RIRNC 2 and 3 were also divergent as shown by the phylogenetic tree (supplementary fig. S5, Supplementary Material online). At the same time, the coding genes for RIRNC 2 and 3 were located in two different genome scaffolds (*rimc 2*: scaffold 4 [939,913 bp] and *rimc 3*: scaffold 82 [423,271 bp]) and there is no evident linkage between these two scaffolds. Two possible scenarios can be proposed: 1) a gene duplication may have occurred after HGT of a gene from an ancestral cyanobacterium, followed by the diversification and translocation of the *rimc* copies; 2) two genes were transferred from a cyanobacterial ancestor to *R. irregulare* by two ancient independent events of HGT.

HGT between Intimately Interacting Species

Events of HGT have been reported in eukaryotic microorganisms which are closely interacting. In the case of diploid commercial wine yeast, *Saccharomyces cerevisiae* EC1118, a region of the genome involved in key wine fermentation function was originated from non-*Saccharomyces* species, *Zygosaccharomyces bailii* which shares habitat (Novo et al. 2009). Another case of HGT has been reported to occur between oomycetes and fungi, which was related with genes of virulence and infection of host plants (Soanes and Richards 2014). Many examples of HGT between endosymbiotic prokaryotes and their hosts have been reported, including *Mollicutes-related endobacteria*, the bacterium which forms endosymbiosis with Glomeromycota and Mucoromycotina (Kuo 2015; Naito et al. 2015; Torres-Cortes et al. 2015). Authors found between 3% and 5% of protein coding genes in *Mollicutes-related endobacterial* genomes, were horizontally transferred from the ancestors of Glomeromycota and Mucoromycotina (Torres-Cortes et al. 2015). This example involving *Mollicutes-related endobacteria* and AMF ancestor is an example of HGT from AMF ancestor to endosymbiotic bacterial ancestor. However, we cannot ignore the possibility of HGTs in opposite direction: from bacteria to AMF. Indeed, in nature, HGT frequently take place in bidirectional way between symbiotic partners (Soucy et al. 2015).

In our study, we found the event of HGT in core enzyme (ribonuclease III) of sRNA regulation which is reported to be linked with symbiosis as well as gene regulation of the

organism (Thiebaut et al. 2014; Yan et al. 2015; Formey et al. 2016). Moreover, the HGT has likely occurred long time ago (with cyanobacterial ancestor) resulting in the occurrence of two additional enzymes of class I ribonuclease III in *R. irregulare* genome. At the same time, class I ribonuclease III homologs have never been reported in other fungi or other eukaryotes. Our finding of RIRNC 2 and RIRNC 3 in *R. irregulare* which both show the closest sequence similarities with cyanobacterial class I ribonuclease III enzymes (fig. 5) could probably be related with an intimate relationship of *R. irregulare* ancestor with cyanobacterial ancestor.

Observation of class I ribonuclease III in *R. irregulare* is unique, considering that the enzyme's distribution has been known to be limited within prokaryotes (Lamontagne et al. 2001; Liang et al. 2014). However, this observation seems not to be restricted to *R. irregulare* within the phylum Glomeromycota. The homologs of class I ribonuclease III (RIRNC 3) were found among various AMF species in phylum Glomeromycota (*Rhizoglossus irregulare*, *Rhizoglossus clarum*, *Rhizoglossus diaphanum*, *Claroideoglossus etunicatum*, *Scutelospora calospora*) (supplementary fig. S4, Supplementary Material online). On the contrary, the two novel homologs (RIRNC 2 and 3) of class I ribonuclease III found in *R. irregulare* do not exist in other fungal groups including *Mucoromycotina*, which has the closest homology of RNAi protein components with AMF. Moreover, our result of RIRNC 3's RIBOC domain sequence comparison and phylogenetic analysis (fig. 5 and supplementary fig. S5, Supplementary Material online) showed a high level of sequence similarities between AMF species in same genus or order as well as minor variation of sequences of AMF species belonging to different orders. The result may reflect the long time evolution of *rimc 3* gene along the diversification of AMF. Indeed, in phylogenetic analysis of class I ribonuclease III (fig. 5), RIRNC 2 and 3 were still separated from the group of sequences from cyanobacteria, implying the event of HGT has not recently occurred (which would result in higher sequence similarity with modern cyanobacterial ribonuclease III).

Based on RIRNC 3 homologs distribution among five AMF species, a hypothetical HGT event that may occur during the evolution of Glomeromycota is inferred and shown in figure 7. In our study, four species from an order, *Glomerales* (*R. irregulare*, *R. clarum*, *R. diaphanum*, and *C. etunicatum*) and one species (*S. calospora*) from another order, *Diversisporales* were searched for RIRNC 3 homologs possession, and all of five investigated species possessed sequences of RIRNC 3 homologs. We could not specify if HGT occurred in the common ancestor of Glomeromycota, because of the limited availability of AMF genomes of other two basal orders (*Paraglomerales* and *Archaeosporales*) in Glomeromycota. However, according to our result, HGT of RIRNC 3 coding gene from cyanobacterial ancestor occurred at the latest before the divergence of orders *Diversisporales* and *Glomerales* from their common ancestor. What would be the

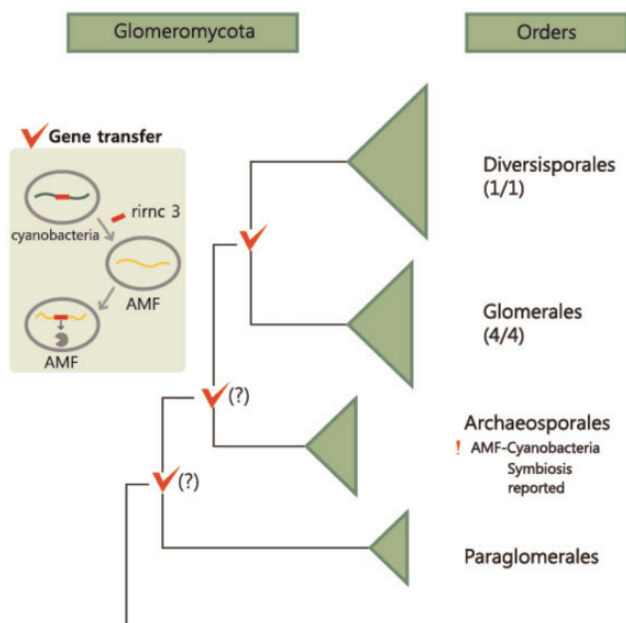


FIG. 7.—Summary of putative HGT of cyanobacterial ribonuclease III (*rirnc 3*) that may occur during the diversification of phylum *Glomeromycota*. In our study, four species from an order, *Glomerales* (*Rhizoglomus irregulare*, *R. clarus* (syn, *R. clarus*), *R. diaphanum* (syn, *R. diaphanous*) and *Claroideoglomus etunicatum*) and one species (*Scutelospora calospora*) from another order, *Diversisporales* were investigated. The numbers indicated below the order names represent (Number of species possesses *rirnc 3* homolog/Number of species genome analyzed). All of 5 different AMF species analyzed possessed cyanobacterial ribonuclease III (*rirnc 3*) homologs. Based on the data set used in our analyses, the event of HGT may have occurred at the latest before the common ancestor of orders *Diversisporales* and *Glomerales* during *Glomeromycota* evolution (indicated by red symbol). Interestingly, AMF-cyanobacteria symbiosis was reported in basal order: *Archaeosporales*.

corresponding knowledge for this unique HGT regarding the conservation of a crucial gene of sRNA regulation in AMF?

Putative Origin of HGT from Cyanobacteria to AMF

As an obligatory symbiont of its host plant, AMF depends for its supplying of carbon nutrition from the photosynthesis of the plant (Harley and Smith 1983). The history of AMF and autotrophic land plant's symbiosis was dated back to at least 460 Ma, according to fossil records and molecular clock analysis (Simon et al. 1993; Redecker et al. 2000). On the other hand, investigations of molecular clocks indicated that the appearance of *Glomeromycota* is much older than the emergence of land plant on earth (Berbee and Taylor 2010; Bidartondo et al. 2011). It was recently reported that fungus-like mycelial fossils were found in a 2.4-billion-year-old basalt (Bengtson et al. 2017), which showed that the fungal lineage is much older and their early evolution may have occurred in the deep ocean. So far, it is uncertain when AMF ancestor lost its ability to assimilate carbon by itself

or if AMF ancestor had former symbiotic partnership with an autotrophic cyanobacterium ancestor. However, the studies of recent molecular data of *Glomeromycota* does not reject the hypothesis that some basal lineages of *Glomeromycota* (at least *Paraglomerales* and *Archaeosporales*) might predate the land plant species (Berbee and Taylor 2010; Bidartondo et al. 2011; Pöggeler and Wöstemeyer 2011).

Interestingly, an evidence of symbiosis between a basal species of *Glomeromycota* (*Geosiphon pyriforme*, assigned to basal order *Archaeosporales*) and Cyanobacteria (*Nostoc punctiforme*) has been documented and characterized (Gehrig et al. 1996). Instead of the photosynthetic land plant as a symbiotic partner, *G. pyriforme* forms symbiosis with photosynthetic cyanobacteria. An analogy can be made between *Geosiphon-Nostoc* symbiosis and *Glomeromycota-Mollicutes-related endobacteria* endosymbiosis where a substantial amount of HGT events were reported (Naito et al. 2015; Torres-Cortes et al. 2015). The *Geosiphon-Nostoc* symbiosis is characterized by the formation of photosynthetically active bladder-like cellular structure that is made of a swollen fungal hypha of up to 2 mm in size, containing *Nostoc* filaments and heterocyst in the upper 2/3 of the bladder and lipid droplets in the lower part (Alexopolous et al. 2004). *Geosiphon* produces large spores resembling those of *Glomeromycota*. Phylogenetic analysis based on SSU rDNA sequence assigned *Geosiphon pyriforme* into order *Archaeosporales* of phylum *Glomeromycota* (Redecker et al. 2013). Compared with other eukaryotes including fungi, the possession of class I ribonuclease III enzyme in AMF is unique and this seems to be related with HGT from cyanobacterial ancestor. It is unclear whether the symbiotic form of *G. pyriforme* is exceptional to this specific species or it is the primitive symbiotic form which was maintained throughout AMF evolution like a living fossil. However, the interaction of *G. pyriforme* and *N. punctiforme* is a clear evidence of proximity of AMF with cyanobacteria. The gene exchange with its symbiotic partner was reported multiple times in *Glomeromycota* (Kuo 2015; Naito et al. 2015; Torres-Cortes et al. 2015). The evolution of RNAi system is related with the symbiosis in eukaryotic organisms including nematode according to many published studies (Yigit et al. 2006; Buck and Blaxter 2013; Weiberg et al. 2015). At the same time, our finding of RNAi system in *R. irregulare* clearly showed its complexity and uniqueness comparing with other fungi, and this could be related to the symbiotic nature with AMF. It is likely that the acquisition of key enzyme in RNAi (cyanobacterial ribonuclease III) by HGT into AMF genomes may reflect an ancient symbiosis history with cyanobacteria. Further analyses of the distribution of cyanobacterial ribonuclease III homologs (RIRNC 2 and 3) in AMF orders: *Archaeosporales* and *Paraglomerales* will bring insights on our hypothesis of HGT event in *Glomeromycota* as well as their complex evolution of symbiosis.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) discovery grant to M.H. (grant number: RGPIN-2017-05067) Discovery fund to M.H. which is gratefully acknowledged. We thank Dr Yves Terrat for his assistance in bioinformatics and Yerim Heo for assistance to build and edit some figures. We also thank three anonymous reviewers for their helpful comments.

Literature Cited

- Alexopolous C, Mims C, Blackwell M. 2004. Introductory mycology, 4th ed. Hoboken (NJ): John Wiley and Sons.
- Altschul SF, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25(17):3389–3402.
- Bengtson S, et al. 2017. Fungus-like mycelial fossils in 2.4-billion-year-old vesicular basalt. *Nat Ecol Evol.* 1(6):0141.
- Berbee ML, Taylor JW. 2010. Dating the molecular clock in fungi – how close are we? *Fungal Biol Rev.* 24(1–2):1–16.
- Bidartondo MI, et al. 2011. The dawn of symbiosis between plants and fungi. *Biol Lett.* 7(4):574–577.
- Bonfante P, Genre A. 2010. Mechanisms underlying beneficial plant-fungus interactions in mycorrhizal symbiosis. *Nat Commun.* 1:48.
- Branscheid A, Devers EA, May P, 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.
- Buck AH, Blaxter M. 2013. Functional diversification of Argonates in nematodes: an expanding universe. *Biochem Soc Trans.* 41(4):881–886.
- Chaib De Mares M, et al. 2015. Horizontal transfer of carbohydrate metabolism genes into ectomycorrhizal *Amanita*. *New Phytol.* 205(4):1552–1564.
- Conrad C, Schmitt JG, Evguenieva-Hackenberg E, Klug G. 2002. One functional subunit is sufficient for catalytic activity and substrate specificity of *Escherichia coli* endoribonuclease III artificial heterodimers. *FEBS Lett.* 518(1–3):93–96.
- Court DL, et al. 2013. RNase III: genetics and function; structure and mechanism. *Annu Rev Genet.* 47:405–431.
- Dang Y, Yang Q, Xue Z, Liu Y. 2011. RNA interference in fungi: pathways, functions, and applications. *Eukaryot Cell* 10(9):1148–1155.
- Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27(8):1164–1165.
- Davis CC, Xi Z. 2015. Horizontal gene transfer in parasitic plants. *Curr Opin Plant Biol.* 26:14–19.
- Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113.
- Formey D, et al. 2016. Regulation of small RNAs and corresponding targets in *Nod* factor-induced *Phaseolus vulgaris* root hair cells. *Int J Mol Sci.* 17(6):887.
- Gehrig H, Schussler A, Kluge M. 1996. *Geosiphon pyriforme*, a fungus forming endocytobiosis with *Nostoc* (cyanobacteria), is an ancestral member of the Glomales: evidence by SSU rRNA analysis. *J Mol Evol.* 43(1):71–81.
- Guindon S, et al. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 59(3):307–321.
- Halary S, et al. 2013. Mating type gene homologues and putative sex pheromone-sensing pathway in arbuscular mycorrhizal fungi, a presumably asexual plant root symbiont. *PLoS One* 8(11):e80729.
- Harley JL, Smith SE. 1983. Mycorrhizal symbiosis. London: Academic Press.
- He D, Fu C-J, Baldauf SL. 2016. Multiple origins of eukaryotic *cox15* suggest horizontal gene transfer from bacteria to jakobid mitochondrial DNA. *Mol Biol Evol.* 33(1):122–133.
- Hijri M, Sanders IR. 2005. Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei. *Nature* 433(7022):160–163.
- Hou Y, Lin S. 2009. Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: gene content estimation for dinoflagellate genomes. *PLoS One* 4(9):e6978.
- Katiyar-Agarwal S, Jin H. 2010. Role of small RNAs in host-microbe interactions. *Annu Rev Phytopathol.* 48:225–246.
- Ketting RF. 2011. The many faces of RNAi. *Dev Cell* 20(2):148–161.
- Knip M, Constantin ME, Thordal-Christensen H. 2014. Trans-kingdom cross-talk: small RNAs on the move. *PLoS Genet.* 10(9):e1004602.
- Kuo C-H. 2015. Scrambled and not-so-tiny genomes of fungal endosymbionts. *Proc Natl Acad Sci U S A.* 112(25):7622–7623.
- Lamontagne B, Larose S, Boulanger J, Elela SA. 2001. The RNase III family: a conserved structure and expanding functions in eukaryotic dsRNA metabolism. *Curr Issues Mol Biol.* 3(4):71–78.
- Lamontagne B, Tremblay A, Abou Elela S. 2000. The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. *Mol Cell Biol.* 20(4):1104–1115.
- Letunic I, Doerks T, Bork P. 2015. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res.* 43:D257–D260.
- Liang Y-H, Lavoie M, Comeau M-A, Abou Elela S, Ji X. 2014. Structure of a eukaryotic RNase III postcleavage complex reveals a double-ruler mechanism for substrate selection. *Mol Cell* 54(3):431–444.
- Lin K, et al. 2014. Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. *PLoS Genet.* 10(1):e1004078.
- Simon L, Bousquet J, Lévesque RC, Lalonde M. 1993. Origin and diversification of endomycorrhizal fungi and coincidence with vascular land plants. *Nature* 363(6424):67–69.
- Marleau J, Dalpe Y, St-Arnaud M, Hijri M. 2011. Spore development and nuclear inheritance in arbuscular mycorrhizal fungi. *BMC Evol Biol.* 11:51.
- Martin WF. 2017. Too much eukaryote LGT. *BioEssays* 39(12):12.
- Meng W, Nicholson AW. 2008. Heterodimer-based analysis of subunit and domain contributions to double-stranded RNA processing by *Escherichia coli* RNase III in vitro. *Biochem J.* 410(1):39–48.
- Nadimi M, Beaudet D, Forget L, Hijri M, Lang BF. 2012. Group I intron-mediated trans-splicing in mitochondria of *Gigaspora rosea* and a robust phylogenetic affiliation of arbuscular mycorrhizal fungi with Mortierellales. *Mol Biol Evol.* 29(9):2199–2210.
- Nadimi M, Daubois L, Hijri M. 2016. Mitochondrial comparative genomics and phylogenetic signal assessment of mtDNA among arbuscular mycorrhizal fungi. *Mol Phylogenet Evol.* 98:74–83.
- Naito M, Morton JB, Pawlowska TE. 2015. Minimal genomes of mycoplasma-related endobacteria are plastic and contain host-derived genes for sustained life within Glomeromycota. *Proc Natl Acad Sci U S A.* 112(25):7791–7796.
- Nicholson AW. 1999. Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol Rev.* 23(3):371–390.
- Nicolas FE, Ruiz-Vazquez RM. 2013. Functional diversity of RNAi-associated sRNAs in fungi. *Int J Mol Sci.* 14:15348–15360.

- Nouri E, Breuillin-Sessoms F, Feller U, Reinhardt D. 2014. Phosphorus and nitrogen regulate arbuscular mycorrhizal symbiosis in *Petunia hybrida*. PLoS One 9(6):e90841.
- Novo M, et al. 2009. Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. Proc Natl Acad Sci U S A. 106(38):16333–16338.
- Oehl F, Sieverding E, Palenzuela J, Ineichen K, Alves da Silva G. 2011. Advances in Glomeromycota taxonomy and classification. IMA Fungus 2(2):191–199.
- Pöggeler S, Wöstemeyer J. 2011. Evolution of fungi and fungal-like organisms. Berlin Heidelberg: Springer-Verlag.
- Pickford AS, Catalanotto C, Cogoni C, Macino G. 2002. Quelling in *Neurospora crassa*. Adv Genet. 46:277–303.
- Pinto-Carbo M, et al. 2016. Evidence of horizontal gene transfer between obligate leaf nodule symbionts. ISME J. 10:2092–2105.
- Redecker D, Kodner R, Graham LE. 2000. Glomalean fungi from the Ordovician. Science 289(5486):1920–1921.
- Redecker D, et al. 2013. An evidence-based consensus for the classification of arbuscular mycorrhizal fungi (Glomeromycota). Mycorrhiza 23(7):515–531.
- Riley R, et al. 2014. Extreme diversification of the mating type-high-mobility group (MATA-HMG) gene family in a plant-associated arbuscular mycorrhizal fungus. New Phytol. 201(1):254–268.
- Smith SE, Read DJ. 2008. Mycorrhizal symbiosis. 3rd ed. New York: Academic Press.
- Soanes D, Richards TA. 2014. Horizontal gene transfer in eukaryotic plant pathogens. Annu Rev Phytopathol. 52:583–614.
- Soucy SM, Huang J, Gogarten JP. 2015. Horizontal gene transfer: building the web of life. Nat Rev Genet. 16(8):472–482.
- Stukenbrock EH, Rosendahl S. 2005. Clonal diversity and population genetic structure of arbuscular mycorrhizal fungi (*Glomus* spp.) studied by multilocus genotyping of single spores. Mol Ecol. 14(3):743–752.
- Sun W, Nicholson AW. 2001. Mechanism of action of *Escherichia coli* ribonuclease III. Stringent chemical requirement for the glutamic acid 117 side chain and Mn²⁺ rescue of the Glu117Asp mutant. Biochemistry 40(16):5102–5110.
- Thiebaut F, et al. 2014. Genome-wide identification of microRNA and siRNA responsive to endophytic beneficial diazotrophic bacteria in maize. BMC Genomics 15:766.
- Tisserant E, et al. 2012. The transcriptome of the arbuscular mycorrhizal fungus *Glomus intraradices* (DAOM 197198) reveals functional tradeoffs in an obligate symbiont. New Phytol. 193(3):755–769.
- Tisserant E, et al. 2013. Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. Proc Natl Acad Sci U S A. 110(50):20117–20122.
- Torres-Cortes G, Ghignone S, Bonfante P, Schussler A. 2015. Mosaic genome of endobacteria in arbuscular mycorrhizal fungi: transkingdom gene transfer in an ancient mycoplasma-fungus association. Proc Natl Acad Sci U S A. 112:7785–7790.
- Torres-Martinez S, Ruiz-Vazquez RM. 2016. RNAi pathways in *Mucor*: a tale of proteins, small RNAs and functional diversity. Fungal Genet Biol. 90:44–52.
- Weiberg A, Bellinger M, Jin H. 2015. Conversations between kingdoms: small RNAs. Curr Opin Biotechnol. 32:207–215.
- Yan Z, et al. 2015. Identification of microRNAs and their mRNA targets during soybean nodule development: functional analysis of the role of miR393j-3p in soybean nodulation. New Phytol. 207(3):748–759.
- Yigit E, et al. 2006. Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell 127:747–757.

Associate editor: Eric Baptiste