Gene expression in *Rhizoglomus irregulare* at two different time points of mycorrhiza establishment in *Helianthus annuus* roots, as revealed by RNA-seq analysis

Alberto Vangelisti[1a], Alessandra Turrini[1a*], Cristiana Sbrana[2a], Luciano Avio[1], Tommaso Giordani[1], Lucia Natali[1], Manuela Giovannetti[1], Andrea Cavallini[1]

[1]Department of Agriculture, Food, and Environment, University of Pisa, Pisa, Italy,

[2]CNR, Institute of Agricultural Biology and Biotechnology UOS Pisa, Pisa, Italy

[a] These three authors contributed equally to the work

* Corresponding author: alessandra.turrini@unipi.it

ORCID IDs

Alessandra Turrini	0000-0002-7186-4418
Cristiana Sbrana	0000-0002-8058-8566
Luciano Avio	0000-0003-2468-3400
Lucia Natali	0000-0003-3179-5910
Manuela Giovannetti	0000-0002-0716-7837

Abstract

Arbuscular mycorrhizal fungi (AMF) play a fundamental role in plant growth and nutrition in natural and agricultural ecosystems. Despite the importance of such symbionts, the different developmental changes occurring during the AMF life cycle have not been fully elucidated at the molecular level. Here, the RNA-seq approach was used to investigate *Rhizoglomus irregulare* specific and common transcripts at two different time points of mycorrhizal establishment in *Helianthus annuus in vivo*. Four days after inoculation, transcripts related to cellular remodeling (actin and tubulin), cellular signaling (calmodulin, serine/threonine protein kinase, 14-3-3 protein and calcium transporting ATPase), lipid metabolism (fatty acid desaturation, steroid hormone and glycerophospholipid biosynthesis), and biosynthetic processes were detected. In addition to such transcripts, 16 days after inoculation, expressed genes linked to binding and catalytic activities, ions (K^+, Ca^{2+}, Fe^{2+}, Zn^{2+}, Mn^{2+},Pi, ammonia), sugar and lipid transport and those involved in vacuolar polyphosphate accumulation were found. Knowledge of transcriptomic changes required for symbiosis establishment and performance is of great importance to understand the functional role of AMF symbionts in food crop nutrition and health, and in plant diversity in natural ecosystems.

Keywords

Rhizoglomus irregulare transcriptome; RNA-seq; mycorrhizal colonization stages; sunflower mycorrhizal symbiosis.

Introduction

Arbuscular mycorrhizal (AM) fungi (AMF; phylum Glomeromycota, Tedersoo et al. 2018) play a fundamental role in the complex networks of biotic interactions which link belowground and aboveground plant and soil communities, and represent key biotic factors in natural and agricultural ecosystems (Wardle et al. 2004). As obligate biotrophs, AMF depend on host plants for the completion of their life cycle and their acquisition of organic compounds, mainly sugars (Helber et al. 2011) and lipids (Bravo et al. 2017; Jiang et al. 2017; Keymer et al. 2017; Luginbuehl et al. 2017). In exchange for organic carbon, AMF provide plants with soil mineral nutrients, such as phosphorus (P), nitrogen (N), sulfur (S) potassium (K), calcium (Ca), copper (Cu) and zinc (Zn), absorbed and translocated by means of a large and highly interconnected extraradical mycelium, which grows in the soil beyond the depletion zone around roots, highly increasing nutrient absorbing surface (Smith and Read 2008; Pepe et al. 2016). In addition to plant nutrition, AMF provide essential ecosystem services, such as plant tolerance to biotic and abiotic stresses, soil structure formation, enhancing plant diversity (Gianinazzi et al. 2010; Thirkell et al. 2017). Moreover, they stimulate the biosynthesis of diverse phytochemicals with health-promoting activities, contributing to the production of healthy foods (Sbrana et al. 2014; Avio et al. 2018).

Despite the importance of AMF in natural and in agricultural ecosystems, the different developmental changes occurring during the AMF life cycle have not been adequately investigated at the molecular level. Actually, while the sequential cascade of molecular events occurring during symbiosis establishment has been deeply studied in host plants (Choi et al. 2018), it has still to be fully disclosed in AMF, due to the complexity of their biology and genetics (Sbrana et al. 2017; Kamel et al. 2017a). A first limit to the study of AMF genetic structure is their obligate biotrophy, as AMF can be propagated only in association with their host plants, in pot cultures *in vivo* or in root organ cultures *in vitro*. Moreover, they are multinucleate organisms (a spore can contain hundreds to thousands of nuclei) developing an aseptate mycelium, where nuclei are shared in a common cytoplasm. That the nuclei may have differing genetic organization (heterokaryotic *vs* homokaryotic hypotheses) is still under debate (Croll et al. 2009; Lin et al. 2014; Ropars and Corradi 2015; Ropars et al. 2016). Notwithstanding, advances in DNA sequencing technology allowed the release of the first genome and transcriptome of *Rhizoglomus irregulare* (syn. *Rhizophagus irregularis*) isolate DAOM197198 (Tisserant et al. 2012, 2013), whose genome assembly and annotation were recently improved (Lin et al. 2014; Chen et al. 2018; Balestrini et al. 2019). The complete genome sequencing of five additional *R. irregulare* isolates enabled the detection of differences among isolates at the genome level (Ropars et al. 2016; Chen et al. 2018). In recent years, genomes and/or transcriptomes of other AMF species, such as *Gigaspora margarita* (Salvioli et al. 2016), *Gigaspora rosea* (Tang et al. 2016) and *Rhizoglomus clarum* (syn. *Rhizophagus clarus*) (Sędzielewska Toro and Brachmann 2016) were sequenced by next generation sequencing (NGS) techniques, allowing wide analyses of AMF biology and genetics. Moreover, using an ultra-low input RNA-seq approach, large-scale transcriptome data were obtained from quiescent spores of eight AMF species (*Acaulospora morrowiae*, *Diversispora versiforme*, *Scutellospora calospora*, *Racocetra castanea*, *Paraglomus brasilianum*, *Ambispora leptoticha*, *Claroideoglomus claroideum* and *Funneliformis mosseae*; Beaudet et al. 2017).

Overall, data from different studies showed that AMF share common traits linked to obligate biotrophy, such as the expression of a reduced number of genes involved in plant cell wall degradation compared with pathogenic fungi (Tisserant et al. 2012, 2013; Tang et al. 2016), the overexpression of genes regulating signal transduction, transport (Kikuchi et al. 2016) and the synthesis of hundreds of small secreted proteins (Kamel et al. 2017b; Sędzielewska Toro and Brachmann 2016). Some studies on AM fungal transcriptomes, based on Illumina high-throughput sequencing,

were performed on different stages of the fungal life cycle, such as quiescent spores, pre-symbiotic, symbiotic and extraradical mycelium (Tisserant et al. 2013; Salvioli et al. 2016; Tang et al. 2016). So far, little is known by the RNA-seq technique about fungal gene expression during the progression of a mycorrhizal symbiosis *in vivo*. The advantage of studying *in vivo* gene expression is to avoid genetic and functional variations in isolates cultured *in vitro*, which may only be ascribed to limited host species diversity, high nutrient availability and absence of associated microbiota (Kokkoris and Hart 2019).

Here, we analyzed the transcriptomes obtained in a previous work (Vangelisti et al. 2018) by using a RNA-seq approach, to detect *R. irregulare* transcripts expressed in sunflower roots grown *in vivo*. In the former study, two time points representing early and late stages of mycorrhizal colonization, were selected, namely 4 days after inoculation, when most of the fungal structures occurring in sunflower root cells were represented by appressoria and entry points, and 16 days after inoculation, when 98% of entry points had developed arbuscular colonization.

Material and methods

Fungal material

The AM fungal isolate *R. irregulare* IMA6 was obtained from pot cultures of *Trifolium alexandrinum* as host plants, maintained under greenhouse conditions in the International Microbial Archive (IMA) collection of the Department of Agriculture Food and Environment, University of Pisa, Pisa, Italy. Pot-culture sandy-loamy soil was wet sieved through a 100 μm mesh sieve in order to collect spores, mycelium and fine colonized roots that were used to establish the symbiosis with the root system of *Helianthus annuus* HA412-HO inbred line (USDA accession number PI 642777). Two hundred grams of soil were sieved in order to inoculate 20 plantlets.

Experimental conditions

Mycorrhizal sunflower plants were obtained as described in Vangelisti et al. (2018). Briefly, plants were cultured in 150-mm Petri plates containing steam-sterilized quartz gravel (2 mm average diameter) and inoculated with *R. irregulare* pot culture soil sieving, containing spores, mycelium and finely cut colonised roots. Plants were maintained in a growth chamber at 24 °C and supplied weekly with 6 ml half strength Hoagland's solution.

In order to collect samples for transcriptome analyses, root staining with Trypan blue using lactic acid instead of phenol (Phillips and Hayman 1970) was performed every 48 hours on three plantlets, allowing the detection of the early stage of mycorrhizal establishment, when 72% of entry points were not associated with arbusculated cortical root cells (4 days after inoculation, hereafter 4 d), and of the late stage of mycorrhizal establishment, when 98% of entry points developed arbuscular colonization (16 days after inoculation, hereafter 16 d) (Vangelisti et al. 2018).

RNA extraction

Three whole root systems of mycorrhizal plantlets were collected 4 and 16 days after inoculation, and separately ground in liquid nitrogen. Total RNA was isolated using the Logemann procedure (Logemann et al. 1987). Purification from genomic DNA was carried out using DNaseI (Roche), and a phenol/chloroform purification protocol, followed by standard precipitation procedures, was applied to isolate RNA. RNA quality was evaluated by using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA) together with spectrophotometric and electrophoretic analyses.

Collection of *R. irregulare* cDNA reads

Overall, six libraries of cDNA from the two stages of mycorrhizal colonization (4 and 16 days after inoculation) were prepared. Library construction is described in Vangelisti et al. (2018). Briefly, messenger RNA was converted to cDNA and each library was constructed with the TruSeq RNA Sample Prep Kit (Illumina), according to the manufacturer's protocol (Illumina Inc., San Diego, CA, USA). Hence single-end read sequences (100 bp) were obtained by Illumina HiSeq 2000. Read quality was checked by FastQC (v. 0.11.3) and improved by using Trimmomatic (Bolger et al. 2014), trimming the reads with the following parameters: crop = 95, headcrop = 10, minlen = 85. Possible rRNA traces were removed from all libraries by mapping against sunflower and *R. irregulare* ribosomal sequences obtained from the NCBI database repository. Ribosomal alignment was performed by CLC genomics workbench (v. 9.5.3; default parameters except length fraction = 0.5, similarity fraction =0.8).

Gene expression analysis of *R. irregulare*

Six transcriptomes from *R. irregulare* isolates were used as references for expression analysis (Chen et al. 2018). In detail, transcript sequences from the six isolates were collected, then CD-hit-est (v. 4.6; Li et al. 2001) was performed in order to keep unique sequences and retaining one sequence from redundant transcripts with over 90% similarity. In this way, we constructed a reference transcriptome with overall 66,336 putative transcripts.

Reads from the six cDNA libraries were aligned with the reference transcriptome using CLC genomics workbench (v. 9.5.3.) with strict parameters: mismatch cost=2, gap open cost=3, length fraction=0.8, similarity fraction=0.8. Counts of aligned reads per transcript were normalized by RPKM (Mortazavi et al. 2008) and expression values were filtered for sequences with RPKM > 10 in at least one library.

Annotation tables for each transcript of *R. irregulare*, comprising Gene Ontology (GO) ID, enzyme codes from Kyoto Encyclopedia of Genes and Genomes (KEGG) and signalP were downloaded from the Joint Genome Institute (JGI), as suggested by Chen et al. (2018). Interproscan (v. 5.38.76.0; El-Gebali et al. 2019) was run on the Pfam database using protein sequences in order to retrieve conserved domains. The Pfam showing the largest number of transcripts and those including the most expressed transcripts are reported in supplementary tables (as number of transcripts) and in Figure 4 (including the averaged RPKM of transcripts belonging to each Pfam), respectively. In addition, sequences were submitted to the database for carbohydrate-active enzyme annotation (dbCAN; Yin et al. 2012) to identify potential transcripts belonging to known carbohydrate-active (CAZy) families.

Results

***R. irregulare* IMA6 sequencing**

Six high quality cDNA libraries, obtained from RNA isolated from mycorrhizal roots colonized by *R. irregulare* at 4 and 16 days after inoculation, were used to select fungal reads from Illumina sequencing which generated 215,959,905 sequence reads, each 100 nt in length, encompassing about 29 GB of sequence data, mostly of plant origin; the total number of reads per library ranged from 18.72 to 63.49 million (Vangelisti et al. 2018).

Fungal reads were retrieved by mapping against the reference transcriptome using stringent parameters, overall average percentage of AM reads ranged from 1.56% to 9.44% at the two stages of mycorrhizal establishment. Overall, 10,157 *R. irregulare* predicted transcripts were expressed. A total of 3,593 and 9,785 transcripts were found 4 d and 16 d after inoculation, respectively. Transcripts expressed only at 4 d were 372, those specific for 16 d were 6564, whereas 3,221 transcripts were shared between the two times. The total number of transcripts with a known function was 6,177 (Online Resource 1).

Analysis of *R. irregulare* IMA6 genes expressed 4 days after inoculation

GO term analysis performed on the fungal genes expressed by *R. irregulare* IMA6 4 d after inoculation showed that the most abundant GO terms within the macro-functional class Biological Process were represented by “metabolic process” (76%),“microtubule-based process” (60%) and “biosynthetic process” (23.3%) (Fig. 1). Each of the terms “establishment of localization” and “regulation of biological and cellular processes” represented about 12% of the transcripts in the class, while the 16 remaining terms included fewer than 10% of transcripts. The terms “binding” and “catalytic activity” within Molecular Function, were the most abundant, representing 85% and 32.9% of transcripts, respectively, followed by “peptide transporter activity” (17.5%) and “methylation” (10.1%), the other terms comprising fewer than 10% of transcripts. “Cell part" and "intracellular part” (25.3 and 21.5% respectively), “respiratory chain” (19.8%), “endomembrane system” (18.1%), “protein serine/threonine phosphatase complex” (15.9%), “intracellular organelle” (12.7%) and “clathrin-coated pit” (11.2%) were those mainly characterizing the macro-class Cellular Component, while all the other 15 terms represented fewer than 10% of transcripts in the class (Fig. 1).

The analysis of protein families (Pfam), activated 4 d after inoculation, showed that most of the Pfam detected included transcripts common to the late stages of colonization. Transcripts coding for cytochrome P450, kinases (Protein kinases and Protein tyrosine kinase) and ubiquitin-conjugating enzymes were among the most abundant at the early stage. Other families of proteins involved in transcription/transduction processes, such as those possessing zinc finger, RRM_1, Ras, Rab, Rac, Ran and Ypt1 domains, included a high number of transcripts at the early stage (Fig. 2, Online Resource 2).

Interestingly, the analysis of signal peptides identified 103 predicted signal peptides (HMM signal peptide probability>0.95) among a total of 3,593 transcripts expressed 4 d after inoculation, among which 8 were specific to this time point.

KEGG analysis, providing information on cellular functions and biological systems, showed that lipid metabolism included the highest number of transcripts. The most abundant were those related to fatty acid degradation (33 transcripts, encoding for 9 enzymes) (Tab. 1). Moreover, a high expression of genes related to steroid hormone biosynthesis, arachidonic acid metabolism and glycerophospholipid metabolism was found. A high number of transcripts were also found related to basic cellular metabolism such as glycolysis/gluconeogenesis, pyruvate metabolism and citrate cycle (TCA) (Tab. 1). The highest occurrence of transcripts in carbohydrate metabolism was related to amino sugar and nucleotide sugar metabolism (34 transcripts encoding for 18 enzymes). Deepening the analysis on carbohydrate-acting enzymes (CAZyme), we found 375 transcripts containing domains found in CAZyme families including 139 glycosyltransferases (GTs), 76 glycosidehydrolases (GHs), 27 carbohydrate esterases (CEs), 96 carbohydrate-binding modules (CBMs) and 37 auxiliary activities (AAs) (Online Resource 1).

Among KEGG transcripts involved in amino acid metabolism, those related to tryptophan metabolism were the most represented, with 26 transcripts encoding for 7 enzymes, followed by transcripts involved in cysteine and methionine metabolism and in alanine, aspartate and glutamate metabolism (Tab. 1).

The most expressed transcripts (RPKM mean, Online Resource 1) specific to the 4 d stage of mycorrhizal establishment were a translation initiation factor 4F, a helicase subunit (*eIF-4A*) (523.7 RPKM mean), followed by three transcripts involved in signaling, a calmodulin (EF-Hand superfamily) (378.3 RPKM mean), a serine/threonine protein kinase (178.7 RPKM mean), and a calcium transporting ATPase (37. 2 RPKM mean). Among transcripts specifically expressed only at the 4 d stage we found a chitin synthase (1 transcript) and transcripts involved in

transcription/transduction (Kelch repeat-containing proteins, BTB/POZ domain-containing proteins Rac1 GTPase effector FRL). Transcripts related to transport were also found, such as ABC transporters (13 transcripts), MFS transporters (8 transcripts), an inorganic phosphate transporter (1 transcript), amino acid transporters (2 transcripts) and ammonia permease (2 transcripts), which were not specific, but shared with the 16 d stage. For transcripts related to N metabolism we found glutamine (2 transcripts) and glutamate synthases (1 transcript), carbanoyl phosphate synthetase (1 transcript), arginine succinate synthase (1 transcript), arginase (1 transcript) and ornithine aminotransferase (1 transcript). Interestingly, we also found transcripts related to sugar metabolism, such as threalose-6P synthase component (2 transcripts) and a glycogen synthase (1 transcript) (Online Resource 1).

Analysis of *R. irregulare* IMA6 genes expressed 16 days after inoculation

Results of the GO term analysis carried out on fungal genes expressed by *R. irregulare* IMA6 16 d after inoculation showed that the terms "metabolic" and "biosynthetic" processes (83.9 and 55.7%, respectively) were the most represented within the macro class Biological Process, while the terms "establishment of localization" and "regulation of biological, metabolic and cellular processes" included a lower number of transcripts (less than 10%) compared with those found 4 d after inoculation (Fig. 1). Similarly to what was observed at the 4 d stage, the most abundant terms within the class Molecular Function were "binding" and "catalytic activity", representing 71.6% and 64.7% of transcripts, respectively, followed by "peptide transporter activity" (17.5%) and "methylation" (10.1%). On the contrary, at the 16 d stage the terms "transferase, hydrolase, ligase, signaling receptor and isomerase activity" were not represented in the class Molecular Function, as they included very low levels of transcripts (below 0.1%). The most represented terms in the class Cellular Component showed a similar transcript distribution compared with those detected 4 d after inoculation, with the addition of the term "membrane bounded organelle" (Fig. 1).

The analysis of protein families (Pfam), activated at the 16 d stage, showed that genes included in Pfam classes such as Protein tyrosine and Protein kinase, zinc fingers, WD and BTB/POZ domains and cytochrome P450 were the most abundant (Fig. 3, Online Resource 2). Signal peptide analysis within transcripts expressed at the 16 d stage showed 167 predicted signal peptides (HMM signal peptide probability > 0.95) among a total of 9,803 transcripts (Online resource 1).

KEGG analysis showed that at 16 d, as observed for 4 d, lipid metabolism included the highest number of transcripts, among which those related to fatty acid degradation (90 transcripts, coding for 18 enzymes) were the most abundant (Tab. 2). Moreover, a high expression of genes related to arachidonic acid metabolism, steroid hormone biosynthesis and linoleic acid metabolism was found. The highest occurrence of transcripts in carbohydrate metabolism was related to amino sugar and nucleotide sugar metabolism and to basic cellular functions such as glycolysis/gluconeogenesis, pyruvate metabolism and citrate cycle (TCA), as found at the 4 d stage. A high number of transcripts was also related to inositol phosphate metabolism (Tab. 2). Further analyses showed 794 transcripts containing domains found in CAZyme families, including 273 GTs, 156 GHs, 67 CEs, 240 CBMs and 58 AAs.

Within the KEGG pathways involved in amino acid metabolism we observed a transcript distribution similar to that found at 4 d stage (Tab. 2).

The most expressed transcripts (RPKM mean) detected at the 16 d stage were those encoding for von Willebrand factor and related coagulation proteins, followed by a splicing factor arginine/serine-rich and a Fe^{2+}/Zn^{2+} regulated transporter (Online Resource 1) Interestingly, at the 16 d stage, transcripts related to ion transport, such as transporters, channels, antiporters and symporters, were widely represented, mainly involving K^+ (45 transcripts), Ca^{2+}

(23 transcripts), Fe^{2+}, Zn^{2+}, Mn^{2+} (17 transcripts), Pi (4 transcripts) and Na^{+}/Pi (3 transcripts). Gene expression of a protein involved in vacuolar polyphosphate accumulation was observed only at 16 d. The number of ABC transporters (26 transcripts) and MFS transporters (20 transcripts), specific to the 16 d stage also were abundant. As to transcripts related to N transport and metabolism we found those encoding for ammonia permease (3 transcripts), glutamine synthase (5 transcripts), arginase (1 transcript), argininosuccinate lyase (1 transcript), carbamoyl-P-synthetase (3 transcripts) and ornithine decarboxylase (4 transcripts). Moreover, sugar transporters (7 transcripts), glycogen (4 transcripts) and trehalose synthases (4 transcripts) were found in relation to sugar and lipid transport and metabolism, while choline (1 transcript), cholesterol (1 transcript) and lipid (1 transcript) transporters together with lysophospholipase (6 transcripts) were detected in relation to lipid transport and metabolism (Online Resource 1).

Comparison of genes expressed 4 and 16 days after inoculation

In order to compare transcripts commonly expressed at the two time points, RPKM values of genes belonging to the same GO term were averaged (Online Resource 3). Several GO terms in the three classes showed different levels of expression between the two stages of mycorrhizal development (Online Resource 3). Among common transcripts, the GO terms associated with the highest RPKM means 4 d after inoculation compared to 16 d were GTPase activity (Molecular Function) and cellular protein modification process (Biological Process). On the contrary, the GO terms associated with the highest RPKM values at the 16 d stage, compared to the 4 d stage , were represented by glutamate-ammonia ligase activity, phosphogluconate dehydrogenase (decarboxylating) activity, phospholipase D activity (Molecular Function), superoxide metabolic process, translation, metal ion transport (Biological Process), signal recognition particle, intracellular and proton-transporting two-sector ATPase complex (Cellular Component). In the Molecular Function class, the terms actin, iron and calcium binding, proton-transporting ATPase activity and transporter activity also showed a large difference in RPKM means between the two mycorrhizal colonization times. Interestingly, the terms related to transport in Biological Process displayed higher expression levels 16 d after inoculation compared with 4 d (Online Resource 3).

Mean RPKM of transcripts commonly expressed at the two mycorrhizal colonization times were computed within the Pfam annotations of the most expressed transcripts. At 4 d, transcripts included in the Pfam ubiquitin family showed a larger mean RPKM than those found at 16 d, and the same trend was detected for transcripts belonging to the Pfam tubulin, enolase, elongation factor, 14-3-3 protein, actin, Ring-H2 zinc finger, HSP70 and ATP synthase (Fig. 4, Online Resource 1). On the contrary, transcripts included in the Pfam sugar transporter and fatty acid desaturase displayed higher mean expression levels at 16 d than at 4 d. Similar levels of means RPKM at the two colonization stages were found for transcripts belonging to the Pfam bZip transcription factor and Ras, Ran, Rab domains (Fig. 4, Online Resource 1).

Discussion

In this work, the expression of the gene repertoire of *R. irregulare* IMA6 colonizing the roots of *H. annuus* at two different times of mycorrhizal establishment was analyzed using the RNA-seq approach. The study allowed us to outline a global picture of the *R. irregulare* transcriptomes during mycorrhizal development, highlighting, at the early stage (4 d after inoculation), a high expression of genes related to cellular remodeling, biosynthetic processes, cellular signaling and lipid metabolism. At the later stage (16 d after inoculation), together with the occurrence of many transcripts related to metabolic and biosynthetic processes, binding and catalytic activities, we detected the expression

of genes linked to lipid metabolism, and a high activation of genes encoding for metal ion transporters, suggesting that the deep metabolic changes and cellular remodeling occurring during colonization are highly regulated at the transcriptional level.

During each step of plant symbiosis establishment, the fungal symbiont undergoes drastic physiological and morphological changes, leading first to the production of an intense hyphal branching, and then to appressoria formation, root colonization and penetration into root cells (early stage of symbiosis). After root penetration, intercellular and intracellular hyphae produce arbuscules, which represent the key structure of the symbiosis, where fungal symbiont and host plant mainly exchange nutrient compounds (late stage of symbiosis) (Sbrana et al. 2017). The fungal transcriptome detected 4 days after inoculation likely represents genes activated in the early stage of the symbiosis, because most of the fungal structures occurring in sunflower root cells were represented by appressoria and entry points. The times selected here are in the range reported in previous studies (5-6 and 14-21 days after inoculation) aimed at monitoring AMF gene expression during mycorrhiza establishment using different molecular methods (Breuninger and Requena 2004; Seddas et al. 2008; Hogekamp and Küster 2013).

In this study, most GO terms detected at 4 d were related to metabolic and microtubule-based processes, in particular tubulin and actin showed the largest differences of RPKM mean levels compared with the 16 d stage. Such data suggest that, during the initial phase of mycorrhizal development, the fungal cells were mainly committed to the activation of biological processes involving significant morphological changes. Cytoskeleton proteins, such as myosin I, alpha- and beta-tubulin and interaptin were found at the appressorium formation stage also by Breuninger and Requena (2004), confirming the activation of a process of cell space reorganization. Recently, genes involved in cytoskeleton and cellular organization have been found to be co-transcribed and positively correlated in the two symbiotic partners 18 weeks after inoculation (Mateus et al. 2019), representing key genes for the genetic reprogramming in the mycorrhizal symbiosis. The expression of fungal cytoskeleton-related genes at 4 and 16 d in sunflower mycorrhizal roots highlights their important function in all phases of symbiosis establishment.

At 4 d, an intense activity of transcription and translation occurs in the fungal cells. We found high RPKM values for transcripts related to the translation initiation factor 4F and the helicase subunit (eIF-4A), a DEAD-box protein involved in RNA metabolism (transcription, mRNA splicing and translation, RNA modification and transport, RNA/protein complex assembly and ribosome biogenesis; Andreou and Klostermeier 2013). Such data are consistent with the reprogramming that host plants and fungal symbionts must undergo for establishing the mutualistic interaction (Hogekamp and Küster 2013). Other transcripts showing high RPKM levels were those encoding calmodulin (EF-Hand superfamily) and serine/threonine protein kinase, both involved in signaling. Also, calcium transporting ATPase transcripts were found specifically expressed at a high level 4 d after inoculation. Such data suggest that genes related to Ca^{2+} signaling are highly activated and are very important at the early stage of mycorrhiza establishment, as observed also in other studies (Breuninger and Requena 2004; Liu et al. 2013). Indeed, the expression of genes coding for a calmodulin, a putative p-type Ca^{2+}-ATPase and a Ca^{2+}-induced Ras inactivator (CAPRI), was detected in *F. mosseae* during appressoria development (Breuninger and Requena 2004), while transcripts for Vcx1-like vacuolar Ca^{2+} ion transporter and endoplasmic reticulum putative p-type Ca^{2+}-ATPase were upregulated at the early stage of *R. intraradices* colonization of *Medicago truncatula* (Liu et al. 2013). Such data confirm that a putative signaling cascade, involving Ca^{2+} at the initial phase of symbiotic establishment exists, even if no data are available on molecules involved in the activation of Ca^{2+}, as second messenger in AMF.

316 Among transcripts involved in signaling cascades, 4 and 16 d after inoculation, we detected the expression of a
317 *14-3-3* gene, encoding a multifunctional chaperone, which may have an important role in the formation of intracellular
318 hyphae at the appressorium stage, as suggested by Breuninger and Requena (2004). Recently, Sun et al. (2018)
319 demonstrated that host-induced gene silencing of two genes encoding 14-3-3 proteins (*Ri14-3-3* and *RiBMH2*) impairs
320 arbuscule formation and inhibits the expression of *PT4* and *MST2*, two plant genes fundamental for normal symbiosis
321 functioning. Such data suggest that 14-3-3 proteins are important for the regulation of gene expression also after the
322 penetration stage (Sun et al. 2018). Many genes involved in transcription/transduction processes, such as those
323 possessing Ras, Rab, Rac, Ran and Ypt1 domains were expressed 4 d after inoculation. A high number of transcripts of
324 Ras related small GTPase, RHO type was detected at both stages of sunflower mycorrhization. The important role of
325 the gene encoding RHO protein during mycorrhizal establishment has been demonstrated by Seddas et al. (2009), who
326 showed that such a fungal gene was not active in mycorrhizal *M. truncatula* mutants lacking some important SYM
327 pathway genes, such as *SYMRK/DMI2* and *CCaMK/DMI3*, involved in plant Ca^{2+}-spiking signal at the stage of
328 appressorium development.
329 Additionally, a high level of transcripts related to lipid metabolism were detected 4 d after inoculation. We
330 found many transcripts involved in fatty acid degradation, steroid hormone biosynthesis and glycerophospholipid
331 metabolism, important in membrane invagination and cell remodelling (Wewer et al. 2014). Interestingly, we also found
332 many transcripts encoding for cytochrome P450 which contributes to diverse metabolic pathways required to
333 accommodate the fungal symbiont within plant cells (Handa et al. 2015). Over 200 genes for cytochrome P450 have
334 been detected in the genome of *R. irregulare* (Tisserant et al. 2013), in particular, some subfamilies
335 (CYP4/CYP19/CYP26) are involved in fatty acid metabolism and steroid biosynthesis. A previous study showed that
336 C16 aliphatic fatty acids (cutin monomers) act as plant signals that promote colonization at the appressorium stage
337 (Wang et al. 2012), whilst in mycorrhizal roots propionyl- and butyryl-carnitines, known to be involved in lipid
338 metabolism, could represent new symbiotic signals. These signals would be able to induce AM fungal gene expression
339 of chitin deacetylases, involved in fungal cell-wall organization during fungal growth and plant colonization (Davis and
340 Bartnicki-Garcia 1984; Kamakura et al. 2002), and gene expression of a putative enzyme of the chitooligosaccharide
341 deacetylase family (Laparre et al. 2014). In our work, transcripts important for cell shaping were detected, such as those
342 encoding glycosidehydrolases, chitinase, chitin synthases, and other transcripts related to chitin metabolism. It has been
343 hypothesized that genes for chitin synthase or chitinase also could be involved in the production of
344 lipochitooligosaccharides, fungal signal molecules triggering plant Ca^{2+} spiking at the early stage of mycorrhizal
345 establishment (Malbreil et al. 2014). The fungal accommodation process in sunflower root cells is coupled with host
346 expression of many transcripts belonging to the GO terms “membrane and “cell wall”, supporting the reciprocal
347 integration of the two symbionts (Vangelisti et al. 2018).
348 The most evident transcriptional activities detected 16 d after inoculation were those related to lipid
349 metabolism and metal ion transporters. A high expression of genes encoding ion transporters is expected in mycorrhizal
350 roots, because AMF contribute to plant nutrition by mineral nutrient uptake (mainly P, N, S, K, Ca, Cu, Zn; Tamayo et
351 al. 2014; Calabrese et al. 2016) through a wide network of extraradical mycelium which translocates nutrients to plant
352 roots where arbuscules release them in exchange for plant-fixed carbohydrates (Kiers et al. 2011) and lipids (Jiang et al.
353 2017; Luginbuehl et al. 2017). In sunflower colonized roots, we found different families of fungal transporters (sugar,
354 amino acid, mineral transporters) and channels at both times of mycorrhizal colonization, with an increased number of
355 transcripts and families of transporters when colonization was well established. Transcripts related to the ATP-binding

cassette (ABC) superfamily were the most abundant transcripts encoding transporters at 4 and 16 d after inoculation. Such transporters exploit the energy released by ATP hydrolysis to translocate diverse substrates across membranes (Rees et al. 2009). Our data are consistent with Tisserant and co-authors (2013), who showed upregulation of ABC transporters in intraradical mycelium. At 16 d, we also found the expression of genes encoding different mineral transporters, mainly K^+, Zn^+, Ca^{2+}, Mg^{2+}, Fe, Pi transporters. Moreover, our data suggest that, at 16 d, Pi transport became more important compared to 4 d (transcripts related to Pi transporters were expressed mainly at 16 d). The more intense activity in P accumulation and translocation at 16 d also was confirmed by the expression of a transcript encoding a protein containing an SPX domain involved in vacuolar polyphosphate accumulation, an important trait that allows AMF to acquire a massive amount of P to be delivered to the host plant (Kikuchi et al. 2014). It has been observed that polyphosphate accumulation matches the uptake of Na^+, K^+, Ca^{2+} and Mg^{2+} in the extraradical mycelium, probably neutralizing the negative charge of the molecule (Kikuchi et al. 2014). In our work, the expression of Ca/Mg channels and Na/Pi symporters only at the 16 d stage might be associated with P accumulation also in intraradical mycelium. An increase of transcripts involved in transport of NO_3^-, NH_4^+, auxin, Pi and water was detected in sunflower root cells, confirming that transport modulation is fundamental during the symbiotic interaction (Vangelisti et al. 2018).

In the *R. irregulare* IMA6 transcriptome, at both mycorrhizal colonization stages, we identified diverse transcripts encoding sugar transporters, essential for the development of a functional symbiosis. It has been demonstrated that the reduced expression of such transporters (i.e. RiMST2) led to malformed arbuscules, impaired mycorrhizal colonization and reduced expression of host plant inorganic phosphate transporter PT4 (Helber et al. 2011). The expression of a gene encoding a sucrose transporter observed at 16 d is interesting because AMF were thought not to be able to cleave and assimilate sucrose (Schubert et al. 2004), the main sugar translocated from shoots to roots through the phloem (Giaquinta 1983). The expression of a sucrose transporter also was found in the quiescent spore (Beaudet et al. 2017), but we did not detect transcripts encoding for invertase activity, involved in the cleavage of sucrose, suggesting that sucrose probably is hydrolyzed by host plant enzymes, confirming previous data (Beaudet et al. 2017; Tang et al. 2016). Of note, sunflower mycorrhizal roots showed over-expression of a bidirectional SWEET16-like sugar transporter, an exporter of sucrose and monosaccharides, which represents a possible key element in plant carbon transfer to the fungal symbiont (Vangelisti et al. 2018).

In the transcriptome of *R. irregulare* IMA6, at 16 d we also found a wide expression of lipid transporters which contribute to the transfer of carbon compounds from the host plant to the symbiont. Interestingly, mycorrhizal sunflower roots showed up-regulation of plant genes related to the GO term “lipid metabolic process” at the 16 d stage of colonization (Vangelisti et al. 2018). Here, *R. irregulare* IMA6 transcripts involved in lipid metabolism were the most abundant both at 4 and 16 d. It has been reported that AMF synthesize a large amount of fatty acids for building membranes and storage lipids, through enzymes such as fatty acid elongase, long-chain oxoacyl-CoA reductase, fatty acid D6 desaturase, which are overexpressed in the intraradical mycelium of *R. irregulare* (Wewer et al. 2014). Consistent with those results, we found transcripts encoding a fatty acid D6 desaturase both 4 and 16 d after inoculation. Interestingly, the FA desaturase family showed very high levels of RPKM means at 16 d. On the contrary, AMF are not able to produce C16 fatty acid chains because they lack the cytoplasmic FAS-I complex which synthesizes palmitic acid (Wewer et al. 2014; Ropars et al. 2016; Salvioli et al. 2016; Tang et al. 2016). The lack of the FAS complex was reported in other fungi, such as the fungal pathogen *Malassezia globosa*, that uses secreted lipases to acquire fatty acids from the host (Xu et al. 2007). We found different transcripts related to lipases (33 transcripts at 16 d and 12 transcripts at 4 d), some of which possibly have similar functions. Consistently with the nutrient role of fatty acids in AMF, we

detected a high number of transcripts encoding enzymes involved in fatty acid degradation which were the most represented at both stages. The expression of a high level of transcripts related to lipases, lipid transporters and enzymes involved in fatty acid degradation can lead to speculation about a model of fungal nutrition based mainly on plant-derived lipids in *R. irregulare*/sunflower mutualistic symbiosis. Indeed, recent studies reported that fatty acyl groups obtained from the host plant may represent, in addition to sugars, a key carbon source, necessary for AMF metabolism (Luginbuehl et al. 2017; Jiang et al. 2017, Wewer et al. 2014).

In conclusion, our study expanded previous plant transcriptomic profiling carried out during the establishment of *R. irregulare*/sunflower symbiotic interaction and highlighted a deep modification of gene expression involved in fungal cellular remodeling and signaling at the early root colonization stage, 4 d after inoculation, whilst a high activation of genes encoding ion transporters and proteins related to lipid metabolism were detected at the later stage, 16 d after inoculation. Indeed, lipids represent key cell components in AMF, comprising up to 58% of dry weight in intraradical vesicles (Jabajihare et al. 1984), up to 47% of hyphal volume in extraradical mycelium (Bago et al. 2002) and up to 95% of spore dry weight (Bécard et al. 1991), so that AMF are considered oleaginous fungi (Roth and Paszkowski 2017). The study of the molecular bases of AM fungal developmental processes occurring at different stages of mycorrhizal colonization is of great importance for the identification of subsets of genes regulating fungal sequential reprogramming required for symbiosis establishment and functioning. Knowledge of the genetic bases of AM symbiosis is essential to understand the role of these agriculturally and ecologically important symbionts in crop nutrition and ecosystem conservation.

Funding

This research was funded by a University of Pisa grant (Fondi di Ateneo).

Conflict of interest

The authors declare that they have no conflict of interest.

References

Andreou AZ, Klostermeier D (2013) The DEAD-box helicase eIF4A: paradigm or the odd one out? RNA Biol 10:19-32. https://doi.org/10.4161/rna.21966

Avio L, Turrini A, Giovannetti M, Sbrana C (2018) Designing the ideotype mycorrhizal symbionts for the production of healthy food. Front Plant Sci 9:1089. https://doi.org/10.3389/fpls.2018.01089

Bago B, Zipfel W, Williams RM, Jun J, Arreola R, Lammers PJ, Pfeffer PE, Shachar-Hill Y (2002) Translocation and utilization of fungal storage lipid in the arbuscular mycorrhizal symbiosis. Plant Physiol 128:108-124. https://doi.org/10.1104/pp.010466

Balestrini R, Rosso LC, Veronico P, Melillo MT, De Luca F, Fanelli E, Colagiero M, Salvioli di Fossalunga A, Ciancio A, Pentimone I (2019) Transcriptomic responses to water deficit and nematode infection in mycorrhizal tomato roots. Front Microbiol 10: 1807. https://doi.org/10.3389/fmicb.2019.01807

Beaudet D, Chen EC, Mathieu S, Yildirir G, Ndikumana S, Dalpé Y, Corradi N (2017) Ultra-low input transcriptomics reveal the spore functional content and phylogenetic affiliations of poorly studied arbuscular mycorrhizal fungi. DNA Res 25:217-227. https://doi.org/10.1093/dnares/dsx051

Bécard G, Doner LW, Rolin DB, Douds DD, Pfeffer PE (1991) Identification and quantification of trehalose in vesicular arbuscular mycorrhizal fungi by in vivo C-13 NMR and HPLC analyses. New Phytol 118:547-552. https://doi.org/10.1111/j.1469-8137.1991.tb00994.x

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Bravo A, Brands M, Wewer V, Dormann P, Harrison MJ (2017) Arbuscular mycorrhiza-specific enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular mycorrhiza. New Phytol 214:1631-1645. https://doi.org/10.1111/nph.14533

Breuninger M, Requena N (2004) Recognition events in AM symbiosis: analysis of fungal gene expression at the early appressorium stage. Fungal Genet Biol 41:794-804. https://doi.org/10.1016/j.fgb.2004.04.002

Calabrese S, Pérez-Tienda J, Ellerbeck M, Arnould C, Chatagnier O, Boller T, Schüßler A, Brachmann A, Wipf D, Ferrol N, Courty PE (2016) GintAMT3—a low-affinity ammonium transporter of the arbuscular mycorrhizal *Rhizophagus irregularis*. Front Plant Sci 7:679. https://doi.org/10.3389/fpls.2016.00679

Chen ECH, Morin E, Beaudet D, Noel J, Yildirir G, Ndikumana S, Charron P, St-Onge C, Giorgi J, Kruger M, Marton T, Ropars J, Grigoriev IV, Hainaut M, Henrissat B, Roux C, Martin F, Corradi N (2018) High intraspecific genome diversity in the model arbuscular mycorrhizal symbiont *Rhizophagus irregularis*. New Phytol 220:1161-1171. doi:10.1111/nph.14989

Choi J, Summers W, Paszkowski U (2018) Mechanisms underlying establishment of arbuscular mycorrhizal symbioses. Annu Rev Phytopathol 56:135-160. https://doi.org/10.1146/annurev-phyto-080516-035521

Croll D, Giovannetti M, Koch AM, Sbrana C, Ehinger M, Lammers PJ, Sanders IR (2009) Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus *Glomus intraradices*. New Phytol 181:924-937. https://doi.org/10.1111/j.1469-8137.2008.02726.x

Davis LL, Bartnicki-Garcia S (1984) The co-ordination of chitosan and chitin synthesis in *Mucor rouxii*. J Gen Microbiol 130:2095-2102. https://doi.org/10.1099/00221287-130-8-2095

El-Gebali S, Mistry J, Potter SC, Finn RD (2019) Pfam and MGnify: using metagenomics to improve the Pfam coverage of microbial sequence space. F1000Research 8

Giaquinta RT (1983) Phloem loading of sucrose. Annu Rev Plant Phys 34:347–387. https://doi.org/10.1146/annurev.pp.34.060183.002023

Gianinazzi S, Gollotte A, Binet MN, van Tuinen D, Redecker D, Wipf D (2010) Agroecology: the key role of arbuscular mycorrhizas in ecosystem services. Mycorrhiza 20:519-530. https://doi.org/10.1007/s00572-010-0333-3

Handa Y, Nishide H, Takeda N, Suzuki Y, Kawaguchi M, Saito K (2015) RNA-seq transcriptional profiling of an arbuscular mycorrhiza provides insights into regulated and coordinated gene expression in *Lotus japonicus* and *Rhizophagus irregularis.* Plant Cell Physiol 56:1490-1511. https://doi.org/10.1093/pcp/pcv071

Helber N, Wippel K, Sauer N, Schaarschmidt S, Hause B, 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. https://doi.org/10.1105/tpc.111.089813

Hogekamp C, Küster H (2013) A roadmap of cell-type specific gene expression during sequential stages of the arbuscular mycorrhiza symbiosis. BMC Genomics 14:306. https://doi.org/10.1186/1471-2164-14-306

Jabajihare S, Deschene A, Kendrick B (1984) Lipid-content and composition of vesicles of a vesicular-arbuscular mycorrhizal fungus. Mycologia 76:1024-1030. https://doi.org/10.1080/00275514.1984.12023946

Jiang Y, Wang W, Xie Q, Liu N, Liu L, Wang D, Zhang X, Yang C, Chen X, Tang D, Wang E (2017) Plants transfer lipids to sustain colonization by mutualistic mycorrhizal and parasitic fungi. Science 356:1172-1175. https://doi.org/10.1126/science.aam9970

Kamakura T, Yamaguchi S, Saitoh K, Teraoka T, Yamagucho I (2002) A novel gene *CBP1* encoding a putative extracellular chitin-binding protein may play an important role in the hydrophobic surface sensing of *Magnaporthe grisea* during appressorium differentiation. Mol Plant Microbe Interact 15:437-444. https://doi.org/10.1094/MPMI.2002.15.5.437

Kamel L, Keller-Pearson M, Roux C, Ané JM (2017a) Biology and evolution of arbuscular mycorrhizal symbiosis in the light of genomics. New Phytol 213:531-536. https://doi.org/10.1111/nph.14263

Kamel L, Tang N, Malbreil M, San Clemente H, Le Marquer M, Roux C, Frei dit Frey N (2017b) 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. https://doi.org/10.3389/fpls.2017.00124

Keymer A, Pimprikar P, Wewer V, Huber C, Brands M, Bucerius SL, Delaux P-M, Klingl V, von Ropenack-Lahaye E, Wang TL, Eisenreich W, Dormann P, Parniske M, Guthjahr C (2017) Lipid transfer from plants to arbuscular mycorrhiza fungi. eLife 6:e29107. https://doi.org/10.7554/eLife.29107

Kiers ET, Duhamel M, Beesetty Y, Mensah JA, Franken O, Verbruggen E, Fellbaum CR, Kowalchuk GA, Hart MM, Bago A, Palmer TM, West SA, Vandenkoornhuyse P, Jansa J, Bücking H (2011) Reciprocal rewards stabilize cooperation in the mycorrhizal symbiosis. Science 333:880-882. https://doi.org/10.1126/science.1208473

Kikuchi Y, Hijikata N, Ohtomo R, Handa Y, Kawaguchi M, Saito K, Masuta C, Ezawa T (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. https://doi.org/10.1111/nph.14016

Kikuchi Y, Hijikata N, Yokoyama K, Ohtomo R, Handa Y, Kawaguchi M, Saito K, Ezawa T (2014) Polyphosphate accumulation is driven by transcriptome alterations that lead to near-synchronous and near-equivalent uptake of inorganic cations in an arbuscular mycorrhizal fungus. New Phytol 204:638-649. https://doi.org/10.1111/nph.12937

Kokkoris V, Hart MM (2019) The role of in vitro cultivation on symbiotic trait and function variation in a single species of arbuscular mycorrhizal fungus. Fungal Biol 123:732-744. https://doi.org/10.1016/j.funbio.2019.06.009

Laparre J, Malbreil M, Letisse F, Portais JC, Roux C, Bécard G, Puech-Pagès V (2014) Combining metabolomics and gene expression analysis reveals that propionyl-and butyryl-carnitines are involved in late stages of arbuscular mycorrhizal symbiosis. Mol Plant 7:554-566. https://doi.org/10.1093/mp/sst136

Li W, Jaroszewski L, Godzik A (2001) Clustering of highly homologous sequences to reduce the size of large protein database. Bioinformatics 17:282-283. https://doi.org/10.1093/bioinformatics/17.3.282

Lin K, Limpens E, Zhang Z, Ivanov S, Saunders DGO, Mu D, Pang E, Cao H, Cha H, Lin T, Zhou Q, Shang Y, Li Y, Sharma T, van Velzen R, de Ruijter N, Aanen DK, Win J, Kamoun S, Bisseling T, Geurts R, Huang S (2014) Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. PLoS Genet 10:e1004078. https://doi.org/10.1371/journal.pgen.1004078

Liu Q, Parsons AJ, Xue H, Jones CS, Rasmussen S (2013) Functional characterization and transcript analysis of an alkaline phosphatase from the arbuscular mycorrhizal fungus *Funneliformis mosseae.* Fungal Genet Biol 54:52-59. https://doi.org/10.1016/j.fgb.2013.02.009

Logemann J, Schell J, Willmitzer L (1987) Improved method for the isolation of RNA from plant tissues. Analyt Biochem 163:16-20. https://doi.org/10.1016/0003-2697(87)90086-8

Luginbuehl LH, Menard GN, Kurup S, Van Erp H, Radhakrishnan GV, Breakspear A, Oldroyd GED, Eastmond PJ (2017) Fatty acids in arbuscular mycorrhizal fungi are synthesized by the host plant. Science 356:1175-1178. https://doi.org/10.1126/science.aan0081

Malbreil M, Tisserant E, Martin F, Roux C (2014) Genomics of arbuscular mycorrhizal fungi: out of the shadows. Adv Bot Res 70:259-290. https://doi.org/10.1016/B978-0-12-397940-7.00009-4

Mateus ID, Masclaux FG, Aletti C, Rojas EC, Savary R, Dupuis C, Sanders IR (2019) Dual RNA-seq reveals large-scale non-conserved genotype× genotype-specific genetic reprograming and molecular crosstalk in the mycorrhizal symbiosis. ISME J 13:1226. https://doi.org/10.1038/s41396-018-0342-3

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNASeq. Nature Methods 5:621-628. https://doi.org/10.1038/nmeth.1226

Pepe A, Giovannetti M, Sbrana C (2016) Different levels of hyphal self incompatibility modulate interconnectedness of mycorrhizal networks in three arbuscular mycorrhizal fungi within the Glomeraceae. Mycorrhiza 26:325-332. https://doi.org/10.1007/s00572-015-0671-2

Phillips JM, Hayman DS (1970) Improved procedures for clearing roots andstaining parasitic and vesicular-arbuscular mycorrhizal fungi for rapid assessment of infection. Trans Br Mycol Soc 55:158-161. https://doi.org/10.1016/S0007-1536(70)80110-3

Rees DC, Johnson E, Lewinson O (2009) ABC transporters: the power to change. Nature Rev Mol Cell Biol 10:218. https://doi.org/10.1038/nrm2646

Ropars J, Corradi N (2015) Homokaryotic vs heterokaryotic mycelium in arbuscular mycorrhizal fungi: different techniques different results? New Phytol 208:638-641. https://doi.org/10.1111/nph.13448

Ropars J, Sędzielewska Toro K, Noel J, Pelin A, Charron P, Farinelli L, Marton T, Krüger M, Fuchs J, Brachmann A, Corradi N (2016) Evidence for the sexual origin of heterokaryosis in arbuscular mycorrhizal fungi. Nat Microbiol 1:16033. https://doi.org/10.1038/nmicrobiol.2016.33

Roth R, Paszkowski U (2017) Plant carbon nourishment of arbuscular mycorrhizal fungi. Curr Opin Plant Biol 39:50-56. https://doi.org/10.1016/j.pbi.2017.05.008

Salvioli A, Ghignone S, Novero M, Navazio L, Venice F, Bagnaresi P, Bonfante P (2016) Symbiosis with an endobacterium increases the fitness of a mycorrhizal fungus raising its bioenergetic potential. ISME J 10:130-144. https://doi.org/10.1038/ismej.2015.91

Sbrana C, Avio L, Giovannetti M (2014) Beneficial mycorrhizal symbionts affecting the production of health-promoting phytochemicals. Electrophoresis 35:1535-1546. https://doi.org/101002/elps201300568

Sbrana C, Turrini A, Giovannetti M (2017) The crosstalk between plants and their arbuscular mycorrhizal symbionts: A mycocentric view. In: Richard D, Joseph S (eds) Biocommunication: Sign-mediated interactions between cells and organisms. World Scientific Publishing Europe, pp 285-308.

Schubert A, Allara P, Morte A (2004) Cleavage of sucrose in roots of soybean (*Glycine max*) colonized by an arbuscular mycorrhizal fungus. New Phytol 161:495-501. JSTOR, www.jstor.org/stable/1514332

Seddas PM, Arias CM, Arnould C, Van Tuinen D, Godfroy O, Benhassou HA, Gianinazzi-Pearson V (2009) Symbiosis-related plant genes modulate molecular responses in an arbuscular mycorrhizal fungus during early root interactions. Mol Plant-Microbe Interact 22:341-351. https://doi.org/10.1094/MPMI-22-3-0341

Seddas PM, Arnould C, Tollot M, Arias CM, Gianinazzi-Pearson V (2008) Spatial monitoring of gene activity in extraradical and intraradical developmental stages of arbuscular mycorrhizal fungi by direct fluorescent in situ RT-PCR. Fungal Genet Biol 45:1155-1165. https://doi.org/10.1016/j.fgb.2008.04.013

Sędzielewska Toro K, Brachmann A (2016) The effector candidate repertoire of the arbuscular mycorrhizal fungus *Rhizophagus clarus*. BMC Genomics 17:101. https://doi.org/10.1186/s12864-016-2422-y

Smith SE, Read DJ (2008) Mycorrhizal symbiosis, 3rd Edn. Academic Press, London.

Sun Z, Song J, Xin XA, Xie X, Zhao B (2018) Arbuscular mycorrhizal fungal 14-3-3 proteins are involved in arbuscule formation and responses to abiotic stresses during AM symbiosis. Front Microbiol 9:91. https://doi.org/10.3389/fmicb.2018.00091

Tamayo E, Gómez-Gallego T, Azcón-Aguilar C, Ferrol N (2014) Genome wide analysis of copper, iron and zinc transporters in the arbuscular mycorrhizal fungus *Rhizophagus irregularis*. Front Plant Sci 5:547. doi:10.3389/fpls.2014.00547

564 Tang N, San Clemente H, Roy S, Bécard G, Zhao B, Roux C (2016) A survey of the gene repertoire of *Gigaspora rosea*
565 unravels conserved features among Glomeromycota for obligate biotrophy. Front Microbiol 7:233.
566 https://doi.org/10.3389/fmicb.2016.00233

567 Tedersoo L, Sánchez-Ramírez S, Urmas Köljalg U, Bahram M, Döring M, Schigel D, May T, Ryberg M, Abarenkov K
568 (2018) High-level classification of the Fungi and a tool for evolutionary ecological analyses. Fung Div 90:135-159.
569 https://doi.org/101007/s13225-018-0401-0

570 Thirkell TJ, Charters MD, Elliott AJ, Sait SM, Field KJ (2017) Are mycorrhizal fungi our sustainable saviours?
571 Considerations for achieving food security. J Ecol 105:921-929. https://doi.org/10.1111/1365-2745.12788

572 Tisserant E, Kohler A, Dozolme-Seddas P, Balestrini R, Benabdellah K, Colard A, Croll D, da Silva C, Gomez SK,
573 Koul R, Ferrol N, Fiorilli V, Formey D, Franken P, Helber N, HijriM, Lanfranco L, Lindquist E, Liu Y,Malbreil M,
574 Morin E, Poulain J, Shapiro H, van Tuinen D, Waschke A, Azcón-Aguilar C, Bécard G, Bonfante P, Harrison MJ,
575 Küster H, Lammers P, Paszkowski U, Requena N, Rensing SA, Roux C, Sanders IR, Shachar-Hill Y, Tuskan G,
576 Young JPW, Gianinazzi-Pearson V, Martin F (2012) The transcriptome of the arbuscular mycorrhizal fungus
577 *Glomus intraradices* (DAOM197198) reveals functional tradeoffs in an obligate symbiont. New Phytol 193:755-
578 769. https://doi.org/10.1111/j.1469-8137.2011.03948.x

579 Tisserant E, Malbreil M, Kuo A, Kohler A, Symeonidi A, Balestrini R, Charron P, Duensing N, Frei dit Frey N,
580 Gianinazzi-Pearson V, Gilbert LB, Handa Y, Herr JR, Hijri M, Koul R, Kawaguchi M, Krajinski F, Lammers PJ,
581 Masclaux FG, Murat C, Morin E, Ndikumana S, Pagni M, Petitpierre D, Requena N, Rosikiewicz P, Riley R, Saito
582 K, San ClementeH, Shapiro H, van Tuinen D, Becard G, Bonfante P, Paszkowski U, Shachar-Hill YY, Tuskan GA,
583 Young JPW, Sanders IR, Henrissat B, Rensing SA, Grigoriev IV, Corradi N, Roux C, Martin F (2013) Genome of
584 an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. Proc Natl Acad Sci USA
585 110:20117-20122. https://doi.org/10.1073/pnas.1313452110

586 Vangelisti A, Natali L, Bernardi R, Sbrana C, Turrini A, Hassani-Pak K, Giordani T (2018) Transcriptome changes
587 induced by arbuscular mycorrhizal fungi in sunflower (*Helianthus annuus* L) roots. Sci Rep 8:4.
588 https://doi.org/10.1038/s41598-017-18445-0

589 Wang ET, Schornack S, Marsh JF, Gobbato E, Schwessinger B, Eastmond P, Schultze M, Kamoun S, Oldroyd GED
590 (2012) A common signaling process that promotes mycorrhizal and oomycete colonization of plants. Curr Biol
591 22:2242-2246. https://doi.org/10.1016/j.cub.2012.09.043

592 Wardle DA, Bardgett RD, Klironomos JN, Setälä H, Van Der Putten WH, Wall DH (2004) Ecological linkages between
593 aboveground and belowground biota. Science 304:1629-1633. https://doi.org/10.1126/science.1094875

594 Wewer V, Brands M, Dörmann P (2014) Fatty acid synthesis and lipid metabolism in the obligate biotrophic fungus
595 *Rhizophagus irregularis* during mycorrhization of *Lotus japonicus*. Plant J 79:398-412.
596 https://doi.org/10.1111/tpj.12566

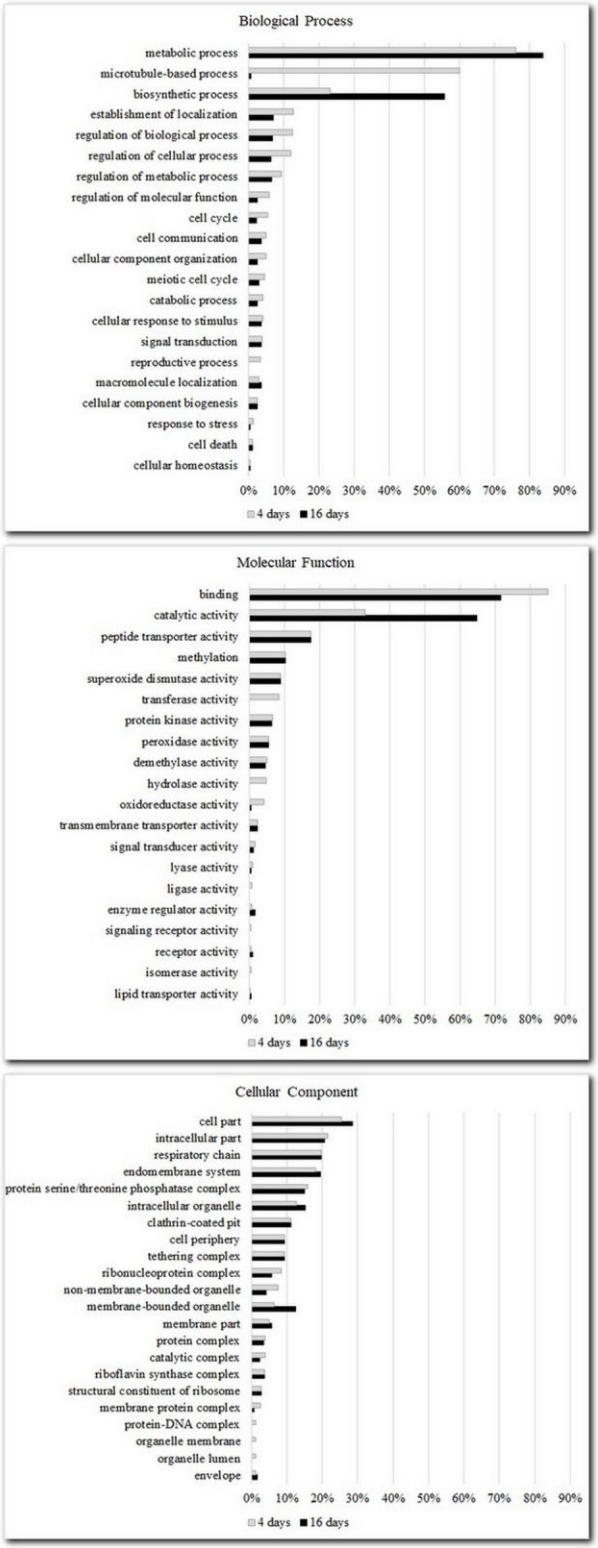
597 Xu J, Saunders CW, Hu P, Grant RA, Boekhout T, Kuramae EE, Kronstad JW, Deangelis YM, Reeder NL, Johnstone
598 KR, Leland M, Fieno AM (2007) Dandruff-associated *Malassezia* genomes reveal convergent and divergent
599 virulence traits shared with plant and human fungal pathogens. Proc Natl Acad Sci USA 104:18730-18735.
600 https://doi.org/10.1073/pnas.0706756104

601 Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y (2012) dbCAN:a web resource for automated carbohydrate-active
602 enzyme annotation. Nucleic Acids Res 40:W445-W451. https://doi.org/101093/nar/gks479

603

604

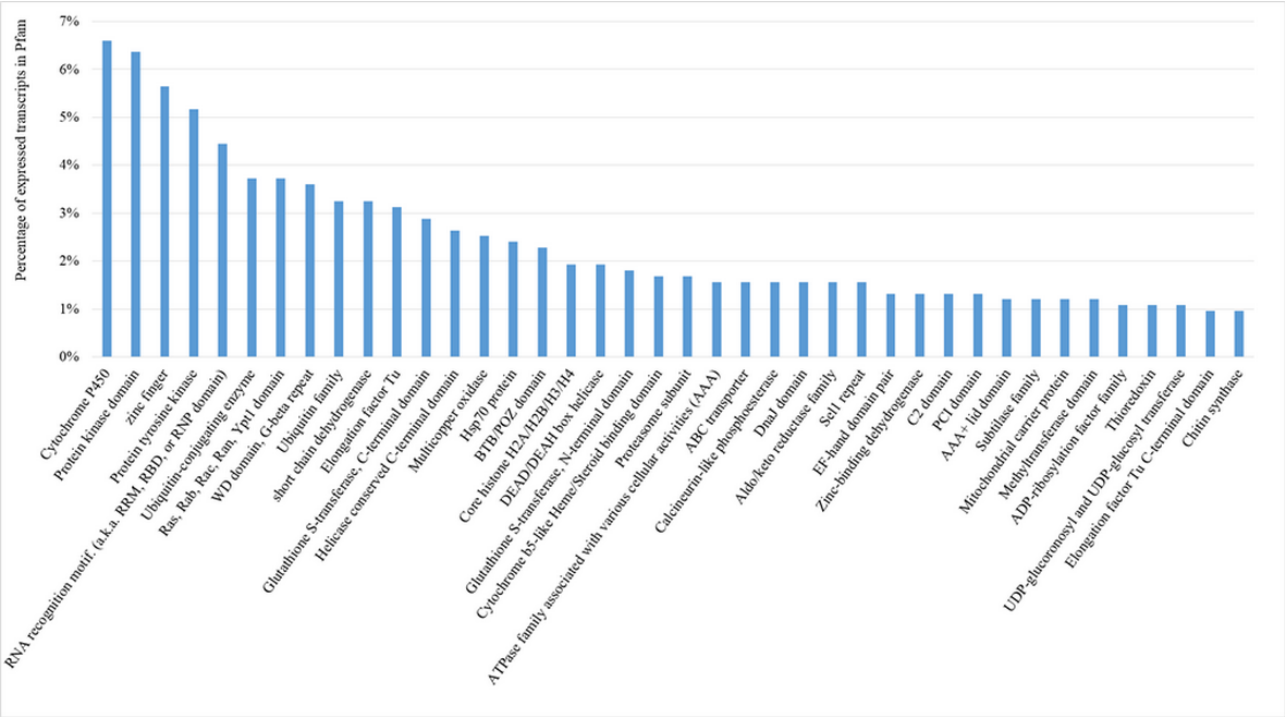
605



606

607 **Figure 1.** GO distribution for *Rhizoglomus irregulare* transcripts obtained from *Helianthus annuus* mycorrhizal roots 4
608 and 16 days after inoculation. The x-axis indicates the percentage of annotated transcripts within each of the three
609 macro-functional classes (Molecular function, Biological process and Cellular component).

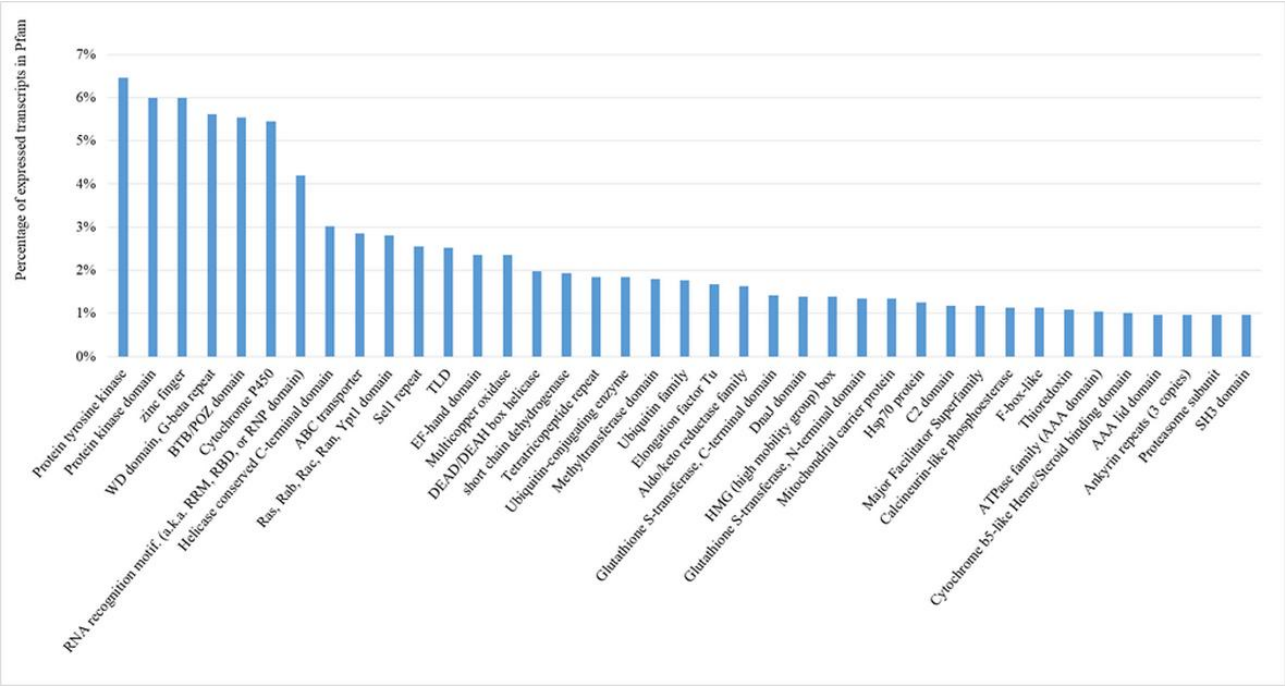
610



611

612 **Figure 2.** Percentages of *Rhizoglomus irregulare* transcripts within the most represented Pfam groups, obtained from
613 *Helianthus annuus* mycorrhizal roots 4 days after inoculation.

614



615

616 **Figure 3.** Percentages of *Rhizoglomus irregulare* transcripts within the most represented Pfam groups, obtained from
617 *Helianthus annuus* mycorrhizal roots 16 days after inoculation.

618

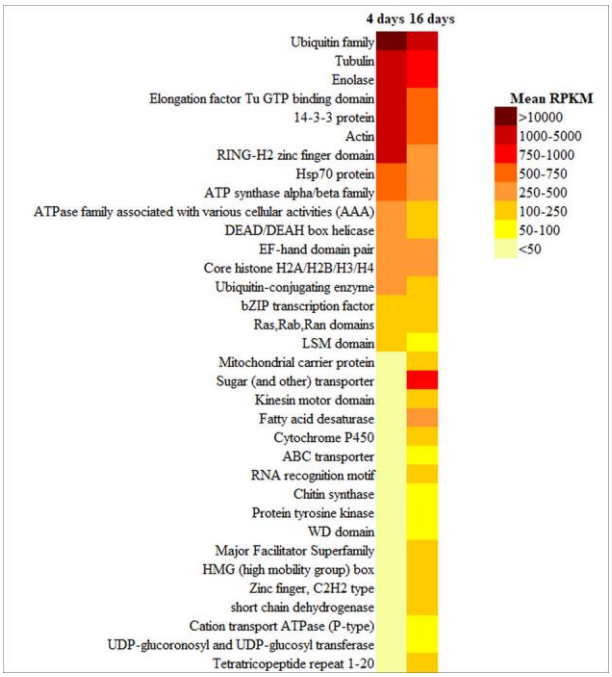


Figure 4. Schematic view of Pfam domains identified in transcripts of *Rhizoglomus irregulare* obtained from *Helianthus annuus* mycorrhizal roots 4 and 16 days after inoculation. RPKM (reads per kilo base per million mapped reads) values of genes belonging to the same Pfam were averaged and used to generate the matrix plot for the two mycorrhizal colonization time points analyzed.

625 **Table 1.** Number of sequences and enzymes within the main KEGG metabolic pathways and annotation of early
626 transcripts of *Rhizoglomus irregulare* obtained from *Helianthus annuus* mycorrhizal roots, 4 days after inoculation.

KEGG pathway	KEGG annotations	sequences	enzymes
Amino acid metabolism	Tryptophan metabolism	26	7
	Cysteine and methionine metabolism	18	13
	Alanine, aspartate and glutamate metabolism	14	12
	Lysine degradation	12	7
	Valine, leucine and isoleucine degradation	11	7
	Arginine biosynthesis	10	8
	Glycine, serine and threonine metabolism	10	10
	Arginine and proline metabolism	9	7
	Phenylalanine metabolism	9	6
	Tyrosine metabolism	7	6
	Phenylalanine, tyrosine and tryptophan biosynthesis	6	5
	beta-Alanine metabolism	5	3
	Histidine metabolism	4	3
	Biotin metabolism	3	3
	Lysine biosynthesis	3	3
	Cyanoamino acid metabolism	1	1
Carbohydrate metabolism	Amino sugar and nucleotide sugar metabolism	34	18
	Citrate cycle (TCA cycle)	21	12
	Pyruvate metabolism	21	13
	Glycolysis / Gluconeogenesis	18	15
	Ascorbate and aldarate metabolism	12	6
	Pentose and glucuronate interconversions	12	6
	Glyoxylate and dicarboxylate metabolism	11	6
	Pentose phosphate pathway	10	10
	Propanoate metabolism	10	8
	Inositol phosphate metabolism	8	5
	Butanoate metabolism	7	3
	Fructose and mannose metabolism	7	6
	Starch and sucrose metabolism	6	6
	Galactose metabolism	3	3
Energy metabolism	Oxidative phosphorylation	34	8
	Photosynthesis	20	2
	Methane metabolism	10	7
	Nitrogen metabolism	7	4
	Sulfur metabolism	1	1
Lipid metabolism	Fatty acid degradation	33	9
	Steroid hormone biosynthesis	27	5
	Arachidonic acid metabolism	26	6
	Glycerophospholipid metabolism	21	13
	Linoleic acid metabolism	15	1
	Ether lipid metabolism	13	4

	Glycerolipid metabolism	9	6
	Sphingolipid metabolism	9	7
	alpha-Linolenic acid metabolism	8	4
	Biosynthesis of unsaturated fatty acids	8	4
	Steroid biosynthesis	8	6
	Fatty acid biosynthesis	5	2
	Fatty acid elongation	5	3
	Steroid degradation	2	2
	Synthesis and degradation of ketone bodies	2	1
Nucleotide metabolism	Purine metabolism	49	22
	Pyrimidine metabolism	25	12
Biosynthesis of other secondary metabolites	Caffeine metabolism	15	1
	Streptomycin biosynthesis	4	3
	Isoquinoline alkaloid biosynthesis	3	2
	Neomycin, kanamycin and gentamicin biosynthesis	3	2
	Tropane, piperidine and pyridine alkaloid biosynthesis	3	2
	Monobactam biosynthesis	2	2
	Novobiocin biosynthesis	1	1
Energy metabolism	Carbon fixation in photosynthetic organisms	16	11
	Carbon fixation pathways in prokaryotes	16	9
Glycan biosynthesis and metabolism	N-Glycan biosynthesis	4	4
	Other glycan degradation	3	1
	Various types of N-glycan biosynthesis	3	3
	Mannose type O-glycan biosynthesis	2	1
	Other types of O-glycan biosynthesis	2	1
	Glycosaminoglycan biosynthesis - heparan sulfate / heparin	1	1
	Glycosphingolipid biosynthesis - lacto and neolacto series	1	1
	Mucin type O-glycan biosynthesis	1	1
Immune system	T cell receptor signaling pathway	17	1
	Th1 and Th2 cell differentiation	17	1
Metabolism of cofactors and vitamins	Retinol metabolism	25	5
	Porphyrin and chlorophyll metabolism	15	8
	Folate biosynthesis	8	6
	Nicotinate and nicotinamide metabolism	6	4
	Vitamin B6 metabolism	5	1
	Riboflavin metabolism	4	4
	Ubiquinone and other terpenoid-quinone biosynthesis	4	3
	One carbon pool by folate	2	2
	Thiamine metabolism	1	1
Metabolism of other amino acids	Glutathione metabolism	29	8
	Selenocompound metabolism	4	4
	Phosphonate and phosphinate metabolism	2	2

	Taurine and hypotaurine metabolism	1	1
Metabolism of terpenoids and polyketides	Terpenoid backbone biosynthesis	6	4
	Geraniol degradation	4	2
	Limonene and pinene degradation	4	2
	Insect hormone biosynthesis	2	1
	Phenylpropanoid biosynthesis	2	1
	Biosynthesis of ansamycins	1	1
	Sesquiterpenoid and triterpenoid biosynthesis	1	1
Signal transduction	Phosphatidylinositol signaling system	7	4
Translation	Aminoacyl-tRNA biosynthesis	16	16
Xenobiotics biodegradation and metabolism	Metabolism of xenobiotics by cytochrome P450	51	7
	Drug metabolism - cytochrome P450	45	5
	Drug metabolism - other enzymes	33	7
	Aminobenzoate degradation	17	2
	Benzoate degradation	6	3
	Caprolactam degradation	4	2
	Toluene degradation	2	1
	Fluorobenzoate degradation	2	1
	Ethylbenzene degradation	2	1
	Chlorocyclohexane and chlorobenzene degradation	2	1
	Chloroalkane and chloroalkene degradation	2	1
	Styrene degradation	1	1

627

628

629 **Table 2.** Number of sequences and enzymes within the main KEGG metabolic pathways and annotation of late
630 transcripts of *Rhizoglomus irregulare* obtained from *Helianthus annuus* mycorrhizal roots, 16 days after inoculation.

KEGG pathway	KEGG annotations	sequences	enzymes
Amino acid metabolism	Tryptophan metabolism	76	17
	Cysteine and methionine metabolism	34	25
	Alanine, aspartate and glutamate metabolism	29	23
	Valine, leucine and isoleucine degradation	29	18
	Lysine degradation	26	16
	Glycine, serine and threonine metabolism	22	19
	Phenylalanine metabolism	20	14
	Arginine and proline metabolism	17	13
	Arginine biosynthesis	16	12
	beta-Alanine metabolism	15	11
	Tyrosine metabolism	15	13
	Phenylalanine, tyrosine and tryptophan biosynthesis	13	10
	Histidine metabolism	11	8
	Lysine biosynthesis	7	7
	Biotin metabolism	4	4
	Cyanoamino acid metabolism	4	3
	D-Glutamine and D-glutamate metabolism	2	1
	Valine, leucine and isoleucine biosynthesis	2	2
Carbohydrate metabolism	Amino sugar and nucleotide sugar metabolism	61	30
	Pyruvate metabolism	33	22
	Citrate cycle (TCA cycle)	29	18
	Glycolysis / Gluconeogenesis	28	23
	Inositol phosphate metabolism	28	13
	Pentose and glucuronate interconversions	25	12
	Glyoxylate and dicarboxylate metabolism	23	14
	Starch and sucrose metabolism	21	15
	Ascorbate and aldarate metabolism	20	8
	Fructose and mannose metabolism	20	15
	Butanoate metabolism	19	12
	Pentose phosphate pathway	18	15
	Propanoate metabolism	18	13
	Galactose metabolism	14	9
Energy metabolism	Oxidative phosphorylation	43	13
	Photosynthesis	21	2
	Methane metabolism	15	11
	Nitrogen metabolism	7	4
	Sulfur metabolism	4	4
Lipid metabolism	Fatty acid degradation	90	18
	Arachidonic acid metabolism	71	12
	Steroid hormone biosynthesis	69	6
	Linoleic acid metabolism	53	2
	Glycerophospholipid metabolism	37	20

	Ether lipid metabolism	33	6
	Glycerolipid metabolism	30	14
	Sphingolipid metabolism	28	13
	Steroid biosynthesis	15	10
	alpha-Linolenic acid metabolism	14	7
	Biosynthesis of unsaturated fatty acids	14	7
	Fatty acid biosynthesis	14	4
	Fatty acid elongation	9	6
	Synthesis and degradation of ketone bodies	5	3
	Steroid degradation	3	3
	Primary bile acid biosynthesis	2	2
Nucleotide metabolism	Purine metabolism	170	49
	Pyrimidine metabolism	94	29
Biosynthesis of other secondary metabolites	Caffeine metabolism	55	3
	Streptomycin biosynthesis	9	7
	Neomycin, kanamycin and gentamicin biosynthesis	5	3
	Isoquinoline alkaloid biosynthesis	4	3
	Tropane, piperidine and pyridine alkaloid biosynthesis	4	3
	Monobactam biosynthesis	3	3
	Novobiocin biosynthesis	2	2
	Acarbose and validamycin biosynthesis	1	1
	Betalain biosynthesis	1	1
	Glucosinolate biosynthesis	1	1
Energy metabolism	Carbon fixation pathways in prokaryotes	29	18
	Carbon fixation in photosynthetic organisms	17	12
Glycan biosynthesis and metabolism	Other glycan degradation	19	4
	N-Glycan biosynthesis	14	10
	Various types of N-glycan biosynthesis	9	7
	Mannose type O-glycan biosynthesis	4	2
	Other types of O-glycan biosynthesis	4	2
	Glycosphingolipid biosynthesis - globo and isoglobo series	3	2
	Glycosaminoglycan biosynthesis - heparan sulfate / heparin	2	2
	Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	2	2
	Glycosaminoglycan degradation	1	1
	Glycosphingolipid biosynthesis - lacto and neolacto series	1	1
	Glycosphingolipid biosynthesis - ganglio series	1	1
	Mucin type O-glycan biosynthesis	1	1
Immune system	T cell receptor signaling pathway	29	2
	Th1 and Th2 cell differentiation	29	2
Metabolism of cofactors and vitamins	Retinol metabolism	69	8
	Porphyrin and chlorophyll metabolism	25	13
	Nicotinate and nicotinamide metabolism	22	12

	Vitamin B6 metabolism	19	3
	Folate biosynthesis	13	11
	One carbon pool by folate	13	11
	Riboflavin metabolism	11	7
	Pantothenate and CoA biosynthesis	8	8
	Ubiquinone and other terpenoid-quinone biosynthesis	6	4
	Thiamine metabolism	2	2
Metabolism of other amino acids	Glutathione metabolism	50	16
	Selenocompound metabolism	8	7
	Taurine and hypotaurine metabolism	4	3
	Phosphonate and phosphinate metabolism	3	3
Metabolism of terpenoids and polyketides	Terpenoid backbone biosynthesis	11	8
	Geraniol degradation	8	5
	Limonene and pinene degradation	8	4
	Phenylpropanoid biosynthesis	5	3
	Insect hormone biosynthesis	4	2
	Biosynthesis of ansamycins	1	1
	Biosynthesis of vancomycin group antibiotics	1	1
	Diterpenoid biosynthesis	1	1
	Polyketide sugar unit biosynthesis	1	1
	Sesquiterpenoid and triterpenoid biosynthesis	1	1
Signal transduction	Phosphatidylinositol signaling system	31	12
Translation	Aminoacyl-tRNA biosynthesis	31	28
Xenobiotics biodegradation and metabolism	Metabolism of xenobiotics by cytochrome P450	103	10
	Drug metabolism - cytochrome P450	97	8
	Aminobenzoate degradation	61	7
	Drug metabolism - other enzymes	53	13
	Benzoate degradation	14	8
	Caprolactam degradation	9	6
	Chloroalkane and chloroalkene degradation	4	2
	Toluene degradation	3	2
	Chlorocyclohexane and chlorobenzene degradation	2	1
	Ethylbenzene degradation	2	1
	Fluorobenzoate degradation	2	1
	Styrene degradation	2	2
	Carbapenem biosynthesis	1	1

631

632