

Rhizophagus Proliferus Genome Sequence Reiterates Conservation of Genetic Traits in AM Fungi, but Predicts Putative Higher Saprotrophic Activity

Pushplata Prasad Singh (✉ pushplata.singh@teri.res.in)

The Energy and Resources institute <https://orcid.org/0000-0002-4273-6898>

Divya Srivastava

The Energy and Resources Institute

Sadhana Shukla

The Energy and Resources Institute

Varsha -

The Energy and Resources Institute

Alok Adholeya

The Energy and Resources Institute

Research Article

Keywords: Rhizophagus proliferus, whole genome sequence, comparative genomics, gene repertoires

Posted Date: March 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-230858/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Archives of Microbiology on December 30th, 2021. See the published version at <https://doi.org/10.1007/s00203-021-02651-6>.

Abstract

Unavailability of the genome sequences of several species of arbuscular mycorrhizal (AM) fungi limits the opportunities for optimizing these biofertilizer species for agricultural benefits. The present work comprises the first draft of the genome sequence of *Rhizophagus proliferus*, which is an important AM species present in biofertilizer consortia for agricultural purpose. The estimated genome size of *Rhizophagus proliferus* is ~ 110 Mbps and the created genomic assembly using the paired-end Illumina reads is 94.35% complete. Genome mining was carried out to identify putative gene families important for biological functions. A total of #22,526 protein-coding genes were estimated in the genome, with an abundance of kinases and reduced number of glycoside hydrolases as compared to other fungal classes. A striking finding in the *R. proliferus* genome was a higher number of carbohydrate esterases (CE), which may suggest towards presence of higher saprotrophic activity in this species as compared to the previously reported AM fungi. The genome sequence and annotation of *R. proliferus* presented here would serve as an important reference for functional genomics studies required for developing biofertilizer formulations in future. In addition, the findings from this work may also prove important in deciphering molecular mechanisms in AM fungi that govern the host-specific interaction and associated agriculture benefits.

1.0 Introduction

AM fungi have approximately 350 to 1000 molecularly defined species under the Division Glomeromycota. These fungi are obligate biotrophs and complete their life-cycle by developing symbiotic relationship with a host plant (Smith and Read 2008). AM fungi have mutualistic association with more than 80% of terrestrial plants (Schüßler et al., 2001; Kivlin et al., 2011) and benefit them by improving nutrient and water uptake efficiency, and also resistance to various abiotic and biotic stresses (Jung et al., 2012). These fungi are endosymbionts and form highly branched-hyphal structures called arbuscules inside the plant cortical cells. The arbuscules deliver mineral nutrients to the cortical cells and also function in carbon acquisition from the host. Mechanisms underlying obligate symbiotic relationship of AM fungi with plant are not completely understood. Knowledge regarding genome organization, genetic functions and reproductive mechanism of several species of AM fungi is not available at present.

Lack of genetic information about gene repertoires of AM fungi poses difficulty in their optimization for crop improvement. Therefore, in order to unravel the genetics of AM fungi, researchers have started exploring genome sequence of different species of AM fungi. Information on gene repertoires of some important species of AM fungi, namely, *Rhizophagus irregularis* (Tisserant et al., 2013), *Rhizophagus clarus* (Kobayashi et al., 2018), *Diversispora epigaea* (Sun et al., 2018), *Gigaspora rosea*, *Rhizophagus diaphanous*, *Rhizophagus cerebriforme* (Morin et al., 2019) and *Gigaspora margarita* (Venice et al., 2020) is now available in the public domain. The most striking findings of these projects include absence of some core eukaryotic genes and presence of a putative sexual reproductive mechanism in AM fungi. An interesting finding reported by researchers involved in understanding the molecular exchanges between AM fungi and the host plant during mycorrhizal interaction is that plant synthesized lipids are imported by AM fungi (Jiang et al., 2017). In agreement with this finding, absence of the multi-domain fatty acid synthase FAS I gene has also been described in the genomes of AM fungi (Tisserant et al., 2013; Luginbueh et al., 2017; Kobayashi et al., 2018; Sun et al., 2018; Morin et al., 2019; Venice et al., 2020).

In the present work, we performed *de novo* genome sequencing and genome annotation of an important AM fungal species, namely, *Rhizophagus proliferus*, and compared the genetic structure with previously reported AM fungi, Ecto-mycorrhizal (EM) fungi and a pathogenic ascomycetous fungus. These included *Rhizophagus irregularis* (Tisserant et al., 2013), *Rhizophagus clarus* (Kobayashi et al., 2018), *Diversispora epigaea* (Sun et al., 2018), *Gigaspora rosea*, *Rhizophagus diaphanous*, *Rhizophagus cerebriforme* (Morin et al., 2019) and *Gigaspora margarita* (Venice et al., 2020) among the AM fungi; *Laccaria bicolor* (Martin et al., 2008) and *Tuber melanosporum* (Martin et al., 2010) among the EM fungi and *Rhizopus oryzae* (Ma et al., 2009) as a pathogenic ascomycetous fungus. *R. proliferus* (previously known as *Glomus proliferum*), is morphologically different from the model AM fungal species *R. irregularis*. *R. proliferus* was first described by Declerck et al., (2000) to possess distinguishing characteristics of spores, such as, small size, hyaline color, smooth wall surface, permanent four-layered spore wall structure, and long hyphae that produced clusters of spores containing several hundred individuals. Most noticeably, anastomoses between hyphae and retraction septa, which are peculiar traits for spore germination in the absence of a host (Logi et al., 1998), were frequently observed in *R. proliferus*. The isolate of *R. proliferus* sequenced in this study has been found to provide agricultural benefits to several crops.

2.0 Material And Methods

2.1 Aim, design and setting of the study

The research work reported here was undertaken to understand the genome structure and function of *R. proliferus*. Genome sequencing was done using Illumina's next-generation sequencing method. A *de novo* assembly was created using the sequenced reads after Quality control. This was followed by *in silico* prediction of gene repertoires in *R. proliferus* followed by their annotation and estimation for important protein families. An additional exploration to identify presence of core eukaryotic genes in *R. proliferus* was also carried out.

2.2 Description of materials

Fungal isolate and DNA extraction

The isolate AM-1901 of *R. proliferus*, from the Centre for Mycorrhizal Culture Collection (CMCC) of The Energy and Resources Institute (TERI), India, was used for genome sequencing in this study. In order to investigate the morphological features of the species microscopic analysis after PVLG and Melzer's staining of spores were carried out using a compound microscope (Carl Zeiss primostar) and a previously published protocol (Błaszowski et al., 2014). Molecular identification of *R. proliferus* was carried out by using SSUmAf-LSUmAr and the SSUmCf-LSUmBr primer pairs in a nested PCR as described by Krüger et al., (2009). Spores were produced in mono-axenic cultures that were maintained on *Agrobacterium rhizogenes*-transformed roots of carrot (*Daucus carota*, Clone GP1). A total of 150,000 sterile spores were collected and high molecular weight (HMW) genomic DNA was extracted.

2.3 Genome Sequencing, assembly and annotation

DNA was fragmented and library was constructed using the Nextera DNA Library Prep protocol. Sequencing (2x150 bp paired-end sequencing) was performed using the services of a commercial service provider (AgriGenome Labs Pvt. Ltd., Kerala, India) on a HiSeq 2500 sequencing platform. Quality control of the DNA library was done by analysis on an Agilent 2000 Bioanalyzer. Preprocessing of reads was carried out (adapter trimming and Q > 20) using AdapterRemoval version 2.2.0. Homology search of preprocessed reads were done using Blastn suite against bacterial database and unaligned reads were considered. The mitochondrial sequences were removed from the bacterial unaligned reads by comparing the reads with the NCBI database mitochondrion.1.1.genomic.fna.gz. The filtered sequences were assembled into scaffolds by executing *De novo* assembly method Spades version 3.12.0 (Bankevich et al., 2012). All scaffolds with length < 1000 bp were excluded in the final assembly. The scaffold sequences were also subjected to homology search in the NCBI nucleotide database to remove non-fungal contamination from carrot-root DNA sequences, all the scaffolds with Identity > 90%, Query-coverage > 75%, GC content > 50% and from non-fungal origin were excluded from the assembly. The repeat sequences were masked using REPEATMASKER version 4.0.7 (<http://www.repeatmasker.org/>) and the total survived scaffolds were considered for downstream analysis. The completeness of the *R. proliferus* draft genome assembly was searched against the core eukaryotic genes present in CEGMA (Parra et al., 2007) to evaluate genome completion. tRNAs were identified using tRNAscan-SE version 1.3.1 (Chan and Lowe 2019). For annotation, gene prediction with taxonomy-based parameters was conducted using AUGUSTUS version 3.1.0 with the gene model of *Saccharomyces cerevisiae* (Stanke et al., 2006). The predicted gene models were annotated by homology search using Diamond version 0.9.3.104 (Buchfink et al., 2015) with NCBI Protein (NR), KEGG (Kanehisa and Goto 2000), InterPro version 5.76.0 (Jones et al., 2014), Pfam (Punta et al., 2012), Gene Ontology (Harris et al., 2004), MATA-HMG (Riley et al., 2014) and MEROPS protease databases (Neil et al., 2018). Prediction for metabolic pathway was carried out by the online KEGG Automatic Annotation Server (KAAS) (Moriya et al., 2007) using the single-directional best-hit information method and default bit score threshold. The annotation of Eukaryotic Orthologous Groups (KOG) and function classes was derived using webMGA tool (Wu et al., 2011).

Gene families for protein kinases and carbohydrate active enzyme (CAZyme) were identified by HMMSCAN searches in Kinomer version 1.0 (Martin et al., 2008(1)) and CAZy databases respectively (Cantarel et al., 2009). G-protein-coupled receptors (GPCRs) in the genome were identified by BLAST search in GPCRdb (Isberg et al., 2016). Homologs of the Pth11-like GPCRs were confirmed by local blastp analysis. Prediction of genes coding for polyketide synthases (PKS), modular nonribosomal peptide synthetases (NRPS), and dimethylallyldiphosphate tryptophan synthases (DMATS) was performed using the SMURF database (Khaldi et al., 2010). Conserved domains were identified using HMMER (Potter et al., 2018) searches against the Pfam database (<https://pfam.xfam.org/>) and SMART (Simple Modular Architecture Research Tool database) (Letunic et al., 2006). Genes coding for the meiosis were additionally searched for homologs in the inventory of 86 genes that were previously reported (Halary et al., 2011). The eukaryotic core genes were identified by blastp analysis (1E-20) against the *Saccharomyces* genome database (Cherry et al., 2012).

3.0 Results

3.1 Morphological and molecular characterization

The morphological details of the isolate of *R. proliferus* sequenced in this study are presented in Fig. 1. The spores were small in size with diameter ranging between 65 to 125 µm at different stages of life-cycle and had three distinct wall layers. The spores were observed to have hyaline color, smooth wall surface, and had long hyphae that held a bunch of spores. Sequencing and phylogenetic analyses of SSU-ITS-LSU nrDNA sequences and morphological studies of spores confirmed the species under investigation to be *R. proliferus*.

3.2 Genome sequencing, assembly and structure

15 Mio reads and 7.8 Gb of primary sequences were received from the whole genome sequencing project. After quality control, the raw sequences were assembled into #12,903 scaffolds with an assembly size of ~ 102.4Mbp and average GC content 27.99%. N₅₀ scaffolds and L₅₀ values were #2126 and #13,544bp respectively (Table 1).

Table 1

Comparative genomic features of *R. proliferus*, *R. irregularis*, *R. clarus*, *R. cerebriforme*, *R. diaphanus*, *G. rosea*, *G. margarita*, *L. bicolor*, *T. melanosporum*, *R. oryzae*

Features	<i>R. proliferus</i>	<i>R. irregularis</i>	<i>R. clarus</i>	<i>R. cerebriforme</i>	<i>R. diaphanus</i>	<i>G. rosea</i>	<i>G. margarita</i>	<i>L. bicolor</i>	<i>T. melanosporum</i>	<i>R. oryzae</i>
Genome Size (Mb)	~ 110 Mb	136.73 Mb	116.42 Mb	136.89 Mb	125.87 Mb	597.95 Mb	773.104 Mb	64.87 Mb	124.94 Mb	39.06 Mb
Number of scaffolds	12,903	1,111	4,424	2,592	2,764	7,526	6,490	665	398	1,168
%GC Content	27.99%	28.7%	27.2	26.55%	27.19%	28.81%	27.68%	47.1%	44.9%	35.4%
N50 scaffold	2126	129	26,609	266	269	734	473	71,011	63,046	3,954
L50	13544	336.38	1,302	147.87	137.49	232.08	-	181	551	2,840
Predicted Proteins	22,526	26,147	27,761	21,549	23,252	31,291	26,603	18,215	7,496	17,467
tRNAs	187	102	-	-	-	-	-	279	143	239
NCBI (blastx) hits	15,087	-	-	-	-	-	-	14,464	5,990	-
Secondary metabolites	0	-	-	-	-	-	-	17	667	22
InterProScan hits	1,670	3,841	-	-	-	-	-	1,432	128	1,556

The assembled *R. proliferus* genome was found #94.35% complete by CEGMA (Table S1). The genomic assembly statistics of *Rhizophagus proliferus*, in comparison to previously reported AM fungi: *Rhizophagus irregularis*, *Rhizophagus cerebriforme*, *Rhizophagus diaphanus*, *Gigaspora rosea*, *Gigaspora margarita*, EM fungi: *Tuber melanosporum*, *Laccaria bicolor*, and pathogenic ascomycetous fungus: *Rhizophus oryzae* are presented in Table 1.

3.3 Genome annotation

A total of #22,526 protein-coding genes were estimated by specifying *Saccharomyces cerevisiae* as the model species in AUGUSTUS version 3.1.0. #187 tRNAs genes were predicted (Table S2). #15,087 proteins shared homology with the NCBI nr database. #3,988 genes that were identified by InterProScan search were found to be distributed in #52 different domains (Table S3). Two predicted domains were unique to *R. proliferus*: PPM-type phosphatase and PTP-type proteins phosphatase, both of which have been found to influence, signal transduction and cell cycle. The terms protein phosphorylation, nitrogen compound metabolic process, cell communication, and signal transduction were frequent GO Biological functions. Hydrolase, transferase and proteins involved in binding of different types of compounds and molecules were among the top 10 terms under GO molecular functions (Fig. 2a, 2b and 2c).

A total of #4569 genes were assigned to #321 KEGG pathways in *R. proliferus* (Table S4) and #1200 genes were predicted under kinases gene families (Kinome) by search in Kinomer version 1.0 (Martin et al., 2008(1)). Proportion of the genes belonging to the kinome family among AM, EM and pathogenic ascomycete fungi was compared (Tisserant et al., 2013; Kobayashi et al., 2018; Sun et al., 2018; Morin et al., 2019; Venice et al., 2020; Martin et al., 2008; Martin et al., 2010; Ma et al., 2009). An expansion of the protein kinase gene family was seen in *R. proliferus* as well as all other AM fungi. Particularly, the Tyrosine kinase-like proteins (TKL), exhibited 20 folds increase in size in *R. proliferus* and other AM species as compared to the ectomycorrhiza fungi *L. bicolor* and around 300 folds increase as compared to the ascomycetes *R. oryzae* (Table 2). Furthermore alpha protein kinases were predicted in *R. proliferus*, which are similar to other AM fungi and are reported to be absent in ectomycorrhizal and pathogenic fungal species.

Table 2
Distribution of protein kinases (kinome) in *R. proliferus* and other mycorrhizal genomes

Genome	AGC	CAMK	CK1	CMGC	STE	TK	TKL	PIKK	PDHK	RIO	Alpha	#N
<i>R. proliferus</i>	23	23	0	5	8	4	1119	2	2	1	13	1200
<i>R. irregularis</i>	25	41	3	24	22	0	832	4	3	2	13	969
<i>L. bicolor</i>	25	38	12	45	18	0	42	3	3	2	0	188
<i>T. melanosporum</i>	19	28	2	23	13	0	1	3	3	2	0	94
<i>R. oryzae</i>	58	93	18	50	55	0	3	5	6	2	0	290

Abbreviations: AGC, protein kinases A, G, and C; CAMK, calcium/calmodulin-dependent kinases; CK1, casein kinase 1; CMGC, cyclin-dependent kinases, mitogen-activated, glycogen-synthase, and CDK-like kinases; STE, sterile phenotype kinases; TK, tyrosine kinases; TKL, tyrosine kinase-like proteins. PIKK, phosphatidylinositol 3-kinase related kinases; PDHK, pyruvate dehydrogenase kinases; RIO, right open reading frame kinases; Alpha, Alpha kinases. Receptor guanylate cyclase kinases (RGC) were not present in any of the species.

#132 genes coding for carbohydrate active enzyme (CAZyme) (Table S5) were identified by HMMSCAN searches in CAZy database. Expansins (EXPN) and polysaccharide lyases (PL) were not predicted in *R. proliferus*. Figure 3 presents the distribution of important classes (AA, CBM, CE, GH, GT, PL and EXPN) under CAZyme in *R. proliferus* in comparison to the previously reported AM, EM and pathogenic ascomycetes fungi (Tisserant et al., 2013; Kobayashi et al., 2018; Sun et al., 2018; Morin et al., 2019; Venice et al., 2020; Martin et al., 2008; Martin et al., 2010; Ma et al., 2009). Interestingly, the Carbohydrate esterases (CE) were present in a significantly higher proportion in *R. proliferus* in comparison to other AM fungi.

G-protein coupled receptors (GPCR) proteins with CFEM domain are involved in different biological processes in fungi such as cellular development and host-pathogen interaction. Only one GPCR-like protein coding gene was identified in *R. proliferus*, but CFEM domain was absent. Query for modular nonribosomal peptide synthetases (NRPS), and dimethylallyl diphosphate tryptophan synthases (DMATS) by using the SMURF database did not yield any gene coding for polyketide synthases (PKS), nonribosomal peptide synthetases (NRPS) or secondary metabolites/toxins.

For genes involved in sexual reproduction, a total of #89 HMG (high mobility group) box containing genes (Table S6) and #47 meiosis-related genes (Table S7) were identified through blastp search in SMART HMG-domain database (Potter et al., 2018). The genes also included the three meiosis-specific genes (Msh4, Dmc1 and Hop2) that have been reported to function exclusively in the meiosis process.

For core eukaryotic genes (CUG) NCBI blastp (1e-20), PFAM (protein family) homology and functional-domain search tools were employed to scan the genome sequence of *R. proliferus*. Distribution of CUG in *R. proliferus* was investigated by blastp analysis (1E-20) against the *Saccharomyces* genome database (Table S8) and #19 genes from the set of "missing ascomycetes core genes (MACGs)" were identified. Furthermore, many important CUG also including the fatty acid synthase (FAS) gene were not found. Table S8 presents a comparative status of presence of these genes in other fungi, which were included for comparison in this study.

4.0 Discussion

AM fungi constitute an important group of fungi for sustainable agriculture benefits; however, the genome sequences and gene repertoires of most of the AM species are not yet explored. The information on genetic structure of these fungi could provide important information about molecular mechanisms underlying the host-specific interaction with different species of crop plants and associated agriculture benefits (Prasad et al., 2019). For majority of fungal classes and species, the information regarding their genetic structure and function has commonly been acquired by comparative studies with the genomes of model species belonging to Ascomycota and Basidiomycota. However, it is difficult to achieve understanding about the genetic architecture of AM fungi by similar comparisons as Ascomycota and Basidiomycota are only distantly related with Glomeromycota and extensive divergence between them over the long evolutionary period has occurred (Sanders and Croll 2010). With such a background, the exceptional identifications regarding the lack of many genes constituting the basic machineries for eukaryotic metabolic pathways in Glomeromycota, expansion of kinome and reduction of CAZymes are being cautiously probed.

The investigation reported here provided first draft of the genome sequence and genome annotation of *R. proliferus*, which is one of the important species of AM fungi known to provide benefits to multiple crops. The estimated size of genome is ~ 110 Mbps, which is the smallest of all the reported AM fungi till date. Like the previously reported AM fungi, conservation with respect to fewer carbohydrate active enzymes and higher number of protein kinases was predicted in *R. proliferus* in comparison to EM and Ascomycetes fungi. High proportion of protein classes representing "establishment of localization" and "signal transduction" proteins were seen in GO classification in *R. proliferus*. Genes coding for "establishment of localization" could be crucially involved in the development of plant-microbial interactions in a symbiotic association.

The remarkable enlargement of protein kinase gene family and especially of tyrosine kinase-like (TKL) genes in is supported by the previous reports in AM fungi (Tisserant et al., 2013; Lin et al., 2014; Salvioli et al., 2016; Tang et al., 2016). Protein kinases influence most cellular activities, especially cell signaling, by protein phosphorylation. The expansion of kinase gene family has been suggested to be crucial in signal transduction

processes that are involved in establishment of symbiotic interaction between AM fungi and plant. Interestingly, TKL-containing proteins have been observed to over-express in germinating spores and intraradical mycelium in *R. irregularis* (Tisserant et al., 2013). Conservation of alpha protein kinases, which is an ancient class of protein (Drennan and Ryazanov 2004), in *R. proliferus* and other AM fungi unlike the other fungal groups, may either indicate inefficiency of AM fungi to expel the genetic load through sexual reproduction or a strong conservation of the molecular mechanisms supportive of lifestyle of AM fungi.

The reduced presence of Glycoside hydrolases found in *R. proliferus* #24 in comparison to other fungal division was in conformity with the other species of AM fungi. Expansins (EXPN) and Polysaccharide Lyases (PL) were absent in *R. proliferus* genome similar to the previous reports in AM species. Expansins (EXPN) functions in cell wall loosening and help the accommodation process of the fungus inside the cortical cells (Cosgrove et al., 2002). Expansins of fungal origin are supposed to function in the loosening of interfacial material loose (Balestrini et al., 2005). Polysaccharides lyases (PL) play a role in degradation of pectin layers of wood (Kristiina and Miia 2018). These observations in AM fungi, unlike the EM and the pathogenic fungi, has been proposed as “functional tradeoffs” in an obligate symbiont for achieving a stealth entry and colonization into root while evading plant immune response (Tisserant et al., 2012). In contrast to the previous reports in AM fungi (Tisserant et al., 2013; Tang et al., 2016; Kobayashi et al., 2018; Sun et al., 2018; Morin et al., 2019; Venice et al., 2020), presence of proteins belonging to AA4 family in *R. proliferus* was striking. AA4 codes for Vanillyl-alcohol oxidase (VAO), which are intracellular FAD-dependent enzymes that act on activated aromatic alcohols like 4-hydroxybenzyl alcohols. AA4 is directly not involved in lignocellulolysis but in the metabolism of lignin-derived compounds. Another noteworthy observation for CAZymes in *R. proliferus* is higher abundance of carbohydrate esterases (CE) in comparison to the other reported AM fungi. CE is a special class of enzyme identified in microorganisms, which de-acetylate hemicellulose and pectin units of plant polysaccharides (Sista and Qin 2018). Deacetylation leads to breaking of glycosidic linkages and help in degradation of plant cell wall, which further enables entry of microorganisms in plants. The abundance of CEs and presence of unique lytic enzyme in *R. proliferus* may indicate towards higher saprotrophic activity in comparison to *R. irregularis*.

A widespread notion of the absence of sexual recombination in AM fungi was challenged by the contrasting observations made in the whole genome analysis of *R. irregularis* (Lin et al., 2014). An exploration of the sexual potential of *R. irregularis* identified a putative AM fungi mating-type locus with prominent similarities to the mating-type locus of Basidiomycota (Ropars et al., 2016). In addition, 76 HMG (high mobility group) box containing genes were identified in *R. irregularis* (Riley et al., 2014). Also, in *G. rosea* #48 meiosis-related genes were found (Tang et al., 2016). In agreement with these findings, #89 HMG (high mobility group) box containing genes and #47 meiosis-related genes were identified in *R. proliferus*. Such a conservation of meiosis-related genes re-emphasized existence of a yet unknown sexual reproduction mechanism in Glomeromycotan fungi and particularly in *R. proliferus*.

In context of presence of CUG in *R. proliferus* a total of 234/248 ultra-conserved CEGs were predicted in *R. proliferus* using CEGMA analysis. The absence of the fatty acid synthesis, type I multienzyme complex (*FAS-I*) was in agreement with previously reported species of AM fungi. Homology search based prediction in *R. proliferus* revealed all components of the bacterial type FAS (type II FAS) genes only. The *FAS-I* complex is responsible for the cytosolic fatty acid synthesis, which produces the bulk of long-chain fatty acids in other fungi (Leibundgut et al., 2008). This gene has been reported missing in the gene repertoires of AM fungi (Tisserant et al., 2013; Wewer et al., 2014; Tang et al., 2016; Sun et al., 2018; Kobayashi et al., 2018; Morin et al., 2019; Venice et al., 2020), which has motivated extensive exploration to understand how AM fungi may generate lipid reserves. Interestingly, in AM-colonized cells of plant roots intensive stimulation of genes involved in lipid metabolism occur, perhaps to provision the increased demand for lipids for the periarbuscular membrane. Based on these findings it has been suggested that AM fungi may receive fatty acids synthesized by plant cells. In this regard, recent studies have demonstrated that AM fungal lipids are, at least partially, derived from the plant host (Bravo et al., 2017; Jiang et al., 2017; Luginbuehl et al., 2017; Rich et al., 2017; Keymer et al., 2018).

Thiamine is a cofactor for enzyme complexes involved in the citric acid cycle, pyruvate dehydrogenase and α -ketoglutarate dehydrogenase, and therefore it is an essential constituent of all cells. The biosynthetic pathway for thiamine has been reported missing in AM fungi. In congruence with the previous reports, thiamine biosynthetic pathway genes were not predicted in *R. proliferus*.

Proteins, uridine permease, uracil permease and dihydroorotate dehydrogenase support uracil metabolism, transport and maintain the intracellular level of uracil. Tight control of the intracellular uracil has been suggested important to reduce the rate of uracil incorporation into DNA (Sun et al., 2013). Dihydroorotate dehydrogenase (DHODH; EC 1.3.99.11), which is the fourth enzyme of the pyrimidine *de novo* biosynthesis pathway, was the only gene from the pathway that was present in both the *R. irregularis* and *R. proliferus* genomes. Genes for glutamate metabolism and glutathione metabolism were predicted in *R. proliferus*, which indicated for its potential for the metabolism of nucleic acids and proteins (Yelamanchi et al., 2016) and detoxification of xenobiotics and the oxidative stress response (Shen et al., 2015) respectively, similar to other AM fungi. Transporters and channels for potassium transport from the soil to the host by the AM fungi are still not completely deciphered in AM fungi. Seven sequences from an EST library of *R. irregularis* were annotated as K^+ transport systems (Casieri et al., 2013), which coded for SKC-type channels and KT/KUP/HAK transporter. Noticeably, no Trk and TOK members were identified in either the EST library or the sequenced nuclear genome (<http://genome.jgi.doe.gov/Gloin1/Gloin1.home.html>). In congruence with the previous reports, no homologue gene for the yeast *TOK1* was identified in *R. proliferus*. However, a Trk-type K^+ transport system in *R. proliferus* was predicted. Probable ferric reductase transmembrane component 8, which is expected to function in the assimilation of iron, was identified only in *R. proliferus* by the conserved domain analysis and

comparison with the proteins coded by yeast. In our comparative analysis, the absence of several CUG in *R. proliferus* was mostly in confirmation with the previously reported AM fungi, which suggested high conservation in genetic features among all species belonging to Glomeromycotina.

5.0 Conclusions

The genome of *R. proliferus* shared several conserved features with previously reported AM species with respect to the genetic structure and functions. This included absence of several eukaryotic genes, prominently the type I FAS gene, abundance of protein kinases and reduced number of glycoside hydrolases. A unique finding was higher proportion of carbohydrate esterases, which might suggest for presence of higher saprotrophic activity in *R. proliferus* as compared to other AM fungi. The present first draft of *R. proliferus* genome would serve as a reference for all future genetics and functional genomics analysis of the species. It would also provide information for comparative genomics analysis required for developing comprehensive understanding about structure and function of AM fungi in future.

Abbreviations

CMCC, Centre for mycorrhizal culture collection

HMW, High molecular weight

NCBI, National Center for Biotechnology Information

KOG, Eukaryotic Orthologous Groups

CAZyme, Carbohydrate active enzyme

BLAST, Basic Local Alignment Search Tool

GPCRs, G-protein-coupled receptors

Blastp, Protein BLAST: search protein databases using a protein query

CDD, NCBI conserved domain database

SMART, Simple Modular Architecture Research Tool database

AA, Auxiliary Activity

CBM, Carbohydrate-Binding Modules

CE, Carbohydrate Esterases

EXPAN, Expansins

GH, Glycoside Hydrolases

GT, Glycosyltransferases

PL, Polysaccharide Lyases

DMAT, DiMethylallyl Tryptophan synthase

NRPS, Non-Ribosomal Peptide Synthetases

PKS, Polyketide Synthetases

TS, Terpene Synthases

Declarations

Funding:

This work was supported by the grant-in-aid for research to Dr. Alok Adholeya by the Department of Biotechnology (DBT), India, under the grant number "BT/PR7075/AGR/21/362/2012". Funders had no role in the study.

Competing interests:

None of the authors declare any competing interests

Availability of data and material:

The whole genome project has been deposited at [https://submit.ncbi.nlm.nih.gov/subs/bioproject/under Bioproject number PRJNA392539](https://submit.ncbi.nlm.nih.gov/subs/bioproject/under/Bioproject%20number/PRJNA392539). All data generated and analysed during this study are included in this report and its supplementary information files.

Authors' contributions

All authors have read and approved the final manuscript. PP was involved in conceptualization of the project, study design, data analyses, data compilation, manuscript writing, critical inputs and finalization of the manuscript. DS was involved in genomics data analysis, data compilation and manuscript writing. Sadhana Shukla carried out DNA extraction and molecular characterization. Varsha was involved in raising monospore cultures and morphology based characterization of spores. AA gave critical comments and reviewed the work.

Ethics approval:

Not Applicable

Consent to participate:

Not Applicable

Consent for publication:

Not Applicable

Acknowledgements

We are thankful to Dr. Reena Singh and Ms. Pratima Vashistha for supporting with spores and morphological details of monospore culture of *Rhizophagus proliferus* isolate sequenced in this study.

References

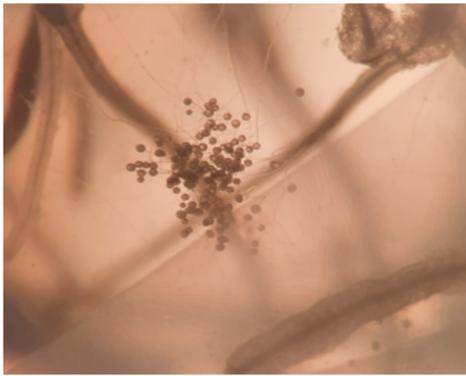
1. Balestrini R, Cosgrove DJ and Bonfante P (2005) Differential location of alpha-expansin proteins during the accommodation of root cells to an arbuscular mycorrhizal fungus. *Planta*, 220(6):889–899. <https://doi.org/10.1007/s00425-004-1431-2>
2. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology: a journal of computational molecular cell biology* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>
3. Błaszczowski J, Chwat G, Góralska A, Ryszka P, Orfanoudakis M (2014) *Septoglomus jasnowskiae* and *Septoglomus turnauae*, two new species of arbuscular mycorrhizal fungi (Glomeromycota). *Mycol Progress* 13:999–1009. <https://doi.org/10.1007/s11557-014-0985-z>
4. Bravo A, Brands M, Wewer V, Dörmann P, Harrison MJ (2017) Arbuscular mycorrhiza-specific enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular mycorrhiza. *New Phytologist* 214:1631–1645. <https://doi.org/10.1111/nph.14533>
5. Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>
6. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* 37:D233–D238. <http://www.cazy.org/>. Accessed 03 January 2020
7. Casieri L, Ait Lahmidi N, Doidy J, Veneault-Fourrey C, Migeon A, Bonneau L, Courty P, Garcia K, Charbonnier M, Delteil A, Brun A, Zimmermann S, Plassard C, Wipf D (2013) Biotrophic transportome in mutualistic plant-fungal interactions. *Mycorrhiza* 23:597–625. <https://doi.org/10.1007/s00572-013-0496-9>

8. Chan PP, Lowe TM (2019) tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. *Methods Mol Biol* 1962:1-14. https://doi.org/10.1007/978-1-4939-9173-0_1
9. Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR, Fisk DG, Hirschman JE, Hitz BC, Karra K, Krieger CJ, Miyasato SR, Nash RS, Park J, Skrzypek MS, Simison M, Weng S, Wong ED (2012) *Saccharomyces Genome Database: the genomics resource of budding yeast*. *Nucleic Acids Res* 40:D700-D705. <https://doi.org/10.1093/nar/gkr1029>
10. Cosgrove DJ, Li LC, Cho HT, Hoffmann-Benning S, Moore RC, Blecker D (2002) The growing world of expansins. *Plant Cell Physiol*. 43:1436-44. <https://doi.org/10.1093/pcp/pcf180>
11. Declerck S, Cranenbrouck S, Dalpé Y, Séguin S, Grandmougin-Ferjani A, Fontaine J, Sancholle M (2000) *Glomus proliferum* sp. nov.: A Description Based on Morphological, Biochemical, Molecular and Monoxenic Cultivation Data. *Mycologia* 92:1178–1187. <https://doi.org/10.1080/00275514.2000.12061265>
12. Drennan D, Ryazanov AG (2004) Alpha-kinases: analysis of the family and comparison with conventional protein kinases. *Prog Biophys Mol Biol* 85:1-32. [https://doi.org/10.1016/S0079-6107\(03\)00060-9](https://doi.org/10.1016/S0079-6107(03)00060-9)
13. Halary S, Malik SB, Lildhar L, Slamovits CH, Hijri M, Corradi N (2011) Conserved meiotic machinery in *Glomus* spp., a putatively ancient asexual fungal lineage. *Genome Biol Evol* 3:950–958. <https://doi.org/10.1093/gbe/evr089>
14. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, Sethuraman A, Theesfeld CL, Botstein D, Dolinski K, Feierbach B, Berardini T, Mundodi S, Rhee SY, Apweiler R, Barrell D, Camon E, Dimmer E, Lee V, Chisholm R, Gaudet P, Kibbe W, Kishore R, Schwarz EM, Sternberg P, Gwinn M, Hannick L, Wortman J, Berriman M, Wood V, de la Cruz N, Tonellato P, Jaiswal P, Seigfried T, White R; Gene Ontology Consortium. (2004) The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*. 32:D258-D261. <https://doi.org/10.1093/nar/gkh036>
15. Isberg V, Mordalski S, Munk C, Rataj K, Harpsøe K, Hauser AS, Vroiling B, Bojarski AJ, Vriend G, Gloriam DE (2016) GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Research* 44:D356-D364. <https://doi.org/10.1093/nar/gkv1178>
16. 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>
17. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236-1240. <https://doi.org/10.1093/bioinformatics/btu031>
18. Jung S C, Martinez-Medina A, Lopez-Raez J A, Pozo M J (2012) Mycorrhiza-Induced Resistance and Priming of Plant Defenses. *Journal of Chemical Ecology* 38:651–664. <https://doi.org/10.1007/s10886-012-0134-6>
19. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 28:27-30. <http://www.genome.ad.jp/kegg/>. Accessed 11 December 2019
20. Keymer A, Huber C, Eisenreich W, Gutjahr C (2018) Tracking Lipid Transfer by Fatty Acid Isotopolog Profiling from Host Plants to Arbuscular Mycorrhiza Fungi. *BIO-PROTOCOL* 8:e2786. <https://doi.org/10.21769/BioProtoc.2786>
21. Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, Fedorova ND (2010) SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genetics Biology*. 47:736-741. www.jcvi.org/smurf. Accessed 06 January 2020
22. Kivlin SN, Hawkes CV, Treseder KK (2011) Global diversity and distribution of arbuscular mycorrhizal fungi. *Soil Biol. Biochem* 43:2294-2303. <https://doi.org/10.1016/j.soilbio.2011.07.012>
23. Kobayashi Y, Maeda T, Yamaguchi K, Kameoka H, Tanaka S, Ezawa T, Shigenobu S, Kawaguchi M (2018) The genome of *Rhizophagus clarus* HR1 reveals a common genetic basis for auxotrophy among arbuscular mycorrhizal fungi. *BMC Genomics* 19(1):465. <https://doi.org/10.1186/s12864-018-4853-0>
24. Kristiina H, Miia Mäkelä R (2018) Role of Fungi in Wood Decay, Reference Module in Life Sciences, Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.12424-0>
25. Krüger M, Stockinger H, Krüger C, Schüßler A (2009) DNA-based species level detection of Glomeromycota: one PCR primer set for all arbuscular mycorrhizal fungi. *New Phytol* 183: 212-223. <https://doi.org/10.1111/j.1469-8137.2009.02835.x>
26. Leibundgut M, Maier T, Jenni S, Ban N (2008) The multienzyme architecture of eukaryotic fatty acid synthases. *Curr Opin Struct Biol* 18:714-25. <https://doi.org/10.1016/j.sbi.2008.09.008>
27. Letunic I, Copley R. R, Pils B, Pinkert S, Schultz J & Bork P (2006) SMART 5: domains in the context of genomes and networks, *Nucleic Acids Research* 34:D257–D260 <http://smart.embl.de/>. Accessed 11 January 2020.
28. Lin K, Limpens E, Zhang Z, Ivanov S, Saunders DG, 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(1):e1004078. <https://doi.org/10.1371/journal.pgen.1004078>

29. Logi C, Sbrana C, Giovannetti M (1998) Cellular events involved in survival of individual arbuscular mycorrhizal symbionts growing in the absence of the host. *Appl Environ Microbiol* 64:3473–3479. <https://doi.org/10.1128/AEM.64.9.3473-3479.1998>
30. 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>
31. Ma LJ, Ibrahim AS, Skory C, Grabherr MG, Burger G, Butler M, Elias M, Idnurm A, Lang BF, Sone T, Abe A, Calvo SE, Corrochano LM, Engels R, Fu J, Hansberg W, Kim JM, Kodira CD, Koehrsen MJ, Liu B, Miranda-Saavedra D, O'Leary S, Ortiz-Castellanos L, Poulter R, Rodriguez-Romero J, Ruiz-Herrera J, Shen YQ, Zeng Q, Galagan J, Birren BW, Cuomo CA, Wickes BL (2009) Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. *PLoS Genet* 5(7):e1000549. <https://doi.org/10.1371/journal.pgen.1000549>
32. Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A, Finn RD (2012) The Pfam protein families database, *Nucleic Acids Research* 40:D1:D290–D301 <https://pfam.xfam.org/>. Accessed 15 December 2019
33. Martin DM, Miranda-Saavedra D, Barton GJ (2008) Kinomer v. 1.0: A database of systematically classified eukaryotic protein kinases. *Nucleic acids research* 37:D244-D250 <http://www.compbio.dundee.ac.uk/kinomer/> . Accessed 02 January 2020
34. Martin F, Aerts A, Ahrén D, Brun A, Danchin EG, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, Salamov A, Shapiro HJ, Wuyts J, Blaudez D, Buée M, Brokstein P, Canbäck B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau A, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, Gay G, Grimwood J, Hoegger PJ, Jain P, Kilaru S, Labbé J, Lin YC, Legué V, Le Tacon F, Marmeisse R, Melayah D, Montanini B, Muratet M, Nehls U, Niculita-Hirzel H, Oudot-Le Secq MP, Peter M, Quesneville H, Rajashekar B, Reich M, Rouhier N, Schmutz J, Yin T, Chalot M, Henrissat B, Kües U, Lucas S, Van de Peer Y, Podila GK, Polle A, Pukkila PJ, Richardson PM, Rouzé P, Sanders IR, Stajich JE, Tunlid A, Tuskan G, Grigoriev IV (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452:88-92. <https://doi.org/10.1038/nature06556>
35. Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM, Jaillon O, Montanini B, Morin E, Noel B, Percudani R, Porcel B, Rubini A, Amicucci A, Amselem J, Anthouard V, Arcioni S, Artiguenave F, Aury JM, Ballario P, Bolchi A, Brenna A, Brun A, Buée M, Cantarel B, Chevalier G, Couloux A, Da Silva C, Denoeud F, Duplessis S, Ghignone S, Hilselberger B, Iotti M, Marçais B, Mello A, Miranda M, Pacioni G, Quesneville H, Riccioni C, Ruotolo R, Splivallo R, Stocchi V, Tisserant E, Viscomi AR, Zambonelli A, Zampieri E, Henrissat B, Lebrun MH, Paolocci F, Bonfante P, Ottonello S, Wincker P (2010) Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* 464:1033–1038. <https://doi.org/10.1038/nature08867>
36. Morin E, Miyauchi S, San Clemente H, Chen ECH, Pelin A, de la Providencia I, Ndikumana S, Beaudet D, Hainaut M, Drula E, Kuo A, Tang N, Roy S, Viala J, Henrissat B, Grigoriev IV, Corradi N, Roux C, Martin FM. (2019) Comparative genomics of *Rhizophagus irregularis*, *R. cerebriforme*, *R. diaphanus* and *Gigaspora rosea* highlights specific genetic features in Glomeromycotina. *New Phytol* 222:1584-1598. <https://doi.org/10.1111/nph.15687>
37. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35:W182-W185 <https://www.genome.jp/kegg/kaas/>. Accessed 27 December 2019
38. Neil D R, Alan J B, Paul D T, Xiaosong H, Alex B, Robert D F (2018) The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research* 46:D624–D632 <https://www.ebi.ac.uk/merops/>. Accessed 24 December 2019
39. Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23:1061–1067. <https://doi.org/10.1093/bioinformatics/btm071>
40. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD (2018) HMMER web server. *Nucleic Acids Research* 46:W200-W204. <https://www.ebi.ac.uk/Tools/hmmer/>. Accessed 08 January 2020
41. Prasad Singh P, Srivastava D, Jaiswar A, Adholeya A (2019) Effector proteins of *Rhizophagus proliferus*: conserved protein domains may play a role in host-specific interaction with different plant species. *Brazilian Journal of Microbiology* 50(3):593-601. <https://doi.org/10.1007/s42770-019-00099-x>
42. Rich MK, Nouri E, Courty PE, Reinhardt D (2017) Diet of Arbuscular Mycorrhizal Fungi: Bread and Butter? *Trends Plant Sci.* 22:652-660. <https://doi.org/10.1016/j.tplants.2017.05.008>
43. Riley R, Charron P, Idnurm A, Farinelli L, Dalpé Y, Martin F, Corradi N (2014) Extreme diversification of the mating type-high-mobility group (MATA-HMG) gene family in a plant-associated arbuscular mycorrhizal fungus. *New Phytol* 201:254-68. <https://doi.org/10.1111/nph.12462>
44. Ropars J, Toro KS, 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. *Nature microbiology* 1(6):16033. <https://doi.org/10.1038/nmicrobiol.2016.33>
45. 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* 10:130-44. <https://doi.org/10.1038/ismej.2015.91>
46. Sanders IR, Croll D (2010) Arbuscular mycorrhiza: the challenge to understand the genetics of the fungal partner. *Annu Rev Genet* 44:271-92. <https://doi.org/10.1146/annurev-genet-102108-134239>

47. Schüßler A, Sschwarzott D, Walker C (2001) A new fungal phylum, the Glomeromycota: phylogeny and evolution. *Mycol. Res.* 105:1413–1421. <https://doi.org/10.1017/S0953756201005196>
48. Shen M, Zhao DK, Qiao Q, Liu L, Wang JL, Cao GH, Li T, Zhao ZW (2015) Identification of glutathione S-transferase (GST) genes from a dark septate endophytic fungus (*Exophiala pisciphila*) and their expression patterns under varied metals stress. *PLoS One* 10(4):e0123418. <https://doi.org/10.1371/journal.pone.0123418>
49. Sista Kameshwar AK, Qin W (2018) Understanding the structural and functional properties of carbohydrate esterases with a special focus on hemicellulose deacetylating acetyl xylan esterases. *Mycology* 9(4):273-295. <https://doi.org/10.1080/21501203.2018.1492979>
50. Smith SE & Read D (2008) *Mycorrhizal Symbiosis*. 3rd edn. Academic press, Australia, pp 611-636
51. Stanke M, Keller O, Gunduz I, Hayes A, Waack S & Morgenstern B (2006) AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Research* 34:W435–W439. <https://doi.org/10.1093/nar/gkl200>
52. Sun X, Chen W, Ivanov S, MacLean A.M, Wight H, Ramaraj T, Mudge J, Harrison M.J. & Fei Z (2018) Genome and evolution of the arbuscular mycorrhizal fungus *Diversispora epigaea* (formerly *Glomus versiforme*) and its bacterial endosymbionts. *New Phytol.* 221:1556-1573. <https://doi.org/10.1111/nph.15472>
53. Sun X, Zhu J, Bao L, Hu C, Jin C, Harris SD, Liu H & Li S (2013) PyrG is required for maintaining stable cellular uracil level and normal sporulation pattern under excess uracil stress in *Aspergillus nidulans*. *Sci China Life Sci* 56:467-75. <https://doi.org/10.1007/s11427-013-4480-6>
54. Tang N, San Clemente H, Roy S, Bécard G, Zhao B & Roux C (2016) A Survey of the Gene Repertoire of *Gigaspora rosea* Unravels Conserved Features among Glomeromycota for Obligate Biotrophy. *Front Microbiol* 7:233. <https://doi.org/10.3389/fmicb.2016.00233>
55. Tisserant E, Kohler A, Dozolme-Seddas P (2012) The transcriptome of the arbuscular mycorrhizal fungus *Glomus intraradices* (DAOM 197198) reveals functional tradeoffs in an obligate symbiont. *New Phytol* 193(3):755-69. <https://doi.org/10.1111/j.1469-8137.2011.03948.x>
56. Tisserant E, Malbreil M, Kuo A et al (2013) Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* 110:20117–20122. <https://doi.org/10.1073/pnas.1313452110>
57. Venice F, Ghignone S, Salvioli di Fossalunga A, Amselem J, Novero M, Xianan X, Sędziewska Toro K, Morin E, Lipzen A, Grigoriev IV, Henrissat B, Martin FM, Bonfante P (2020) “At the nexus of three kingdoms: the genome of the mycorrhizal fungus *Gigaspora margarita* provides insights into plant, endobacterial and fungal interactions.” *Environmental microbiology* 22: 122-141. <https://doi.org/10.1111/1462-2920.14827>
58. Wewer V, Brands M, Dörmann P (2014) Fatty acid synthesis and lipid metabolism in the obligate biotrophic fungus *Rhizophagus irregularis* during mycorrhization of *Lotus japonicus*. *Plant J* 79:398-412. <https://doi.org/10.1111/tpj.12566>
59. Wu S, Zhu Z, Fu L, Niu B, Li W (2011) WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* 12:444 <http://weizhong-lab.ucsd.edu/webMGA/server/kog/> . Accessed 30 December 2019.
60. Yelamanchi SD, Jayaram S, Thomas JK, Gundimeda S, Khan AA, Singhal A, Keshava Prasad TS, Pandey A, Somani BL, Gowda H (2016) A pathway map of glutamate metabolism. *J Cell Commun Signal* 10(1):69-75. <https://doi.org/10.1007/s12079-015-0315-5>

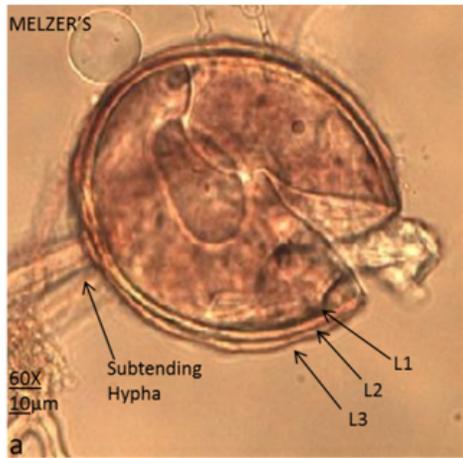
Figures



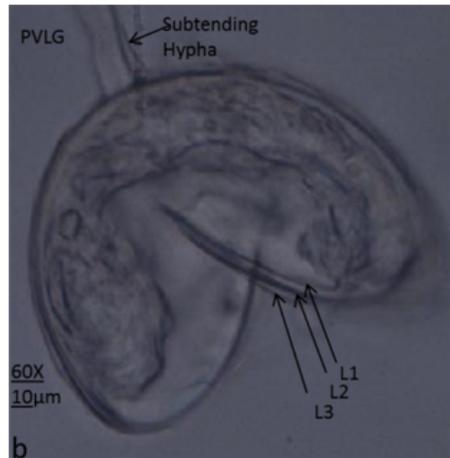
a



b



c



d

Figure 1

Morphology of *R. proliferus*. (a) in vitro culture observed under light inverted microscope at 20 X magnification. (b) spore with subtending hypha observed under compound microscope at 40X magnification. (c) spore morphology observed after Melzer's staining under compound microscope at 40X magnification. (d) spore morphology observed after PVLG staining o under compound microscope at 40X magnification.

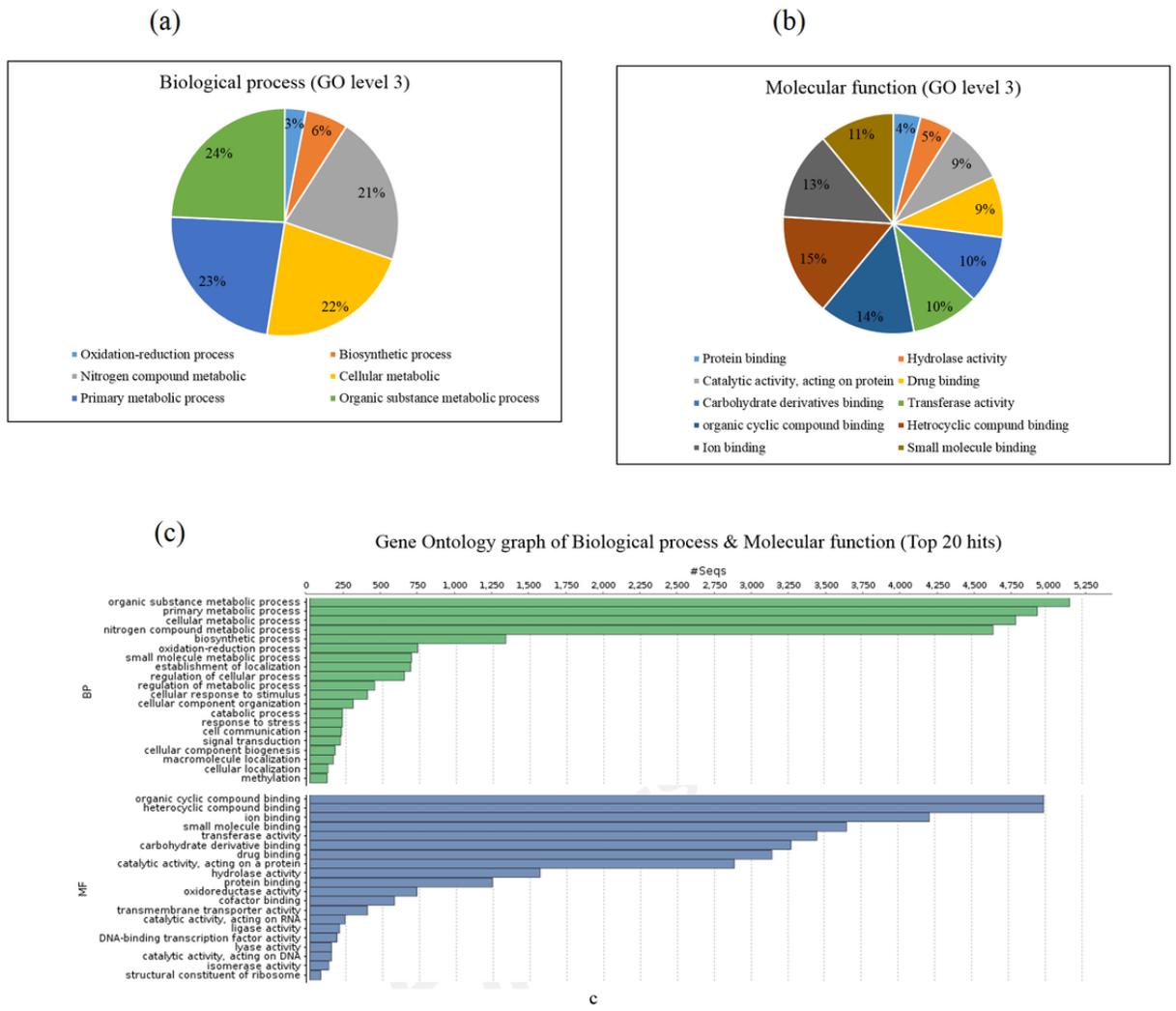


Figure 2

GO-based functional annotation of genes present in the *R. proliferus* genome. (a) Biological process pie chart of GO level 3. (b) Molecular function pie chart of GO level 3. (c) GO classification of Top 20 hits of Biological process & Molecular function.

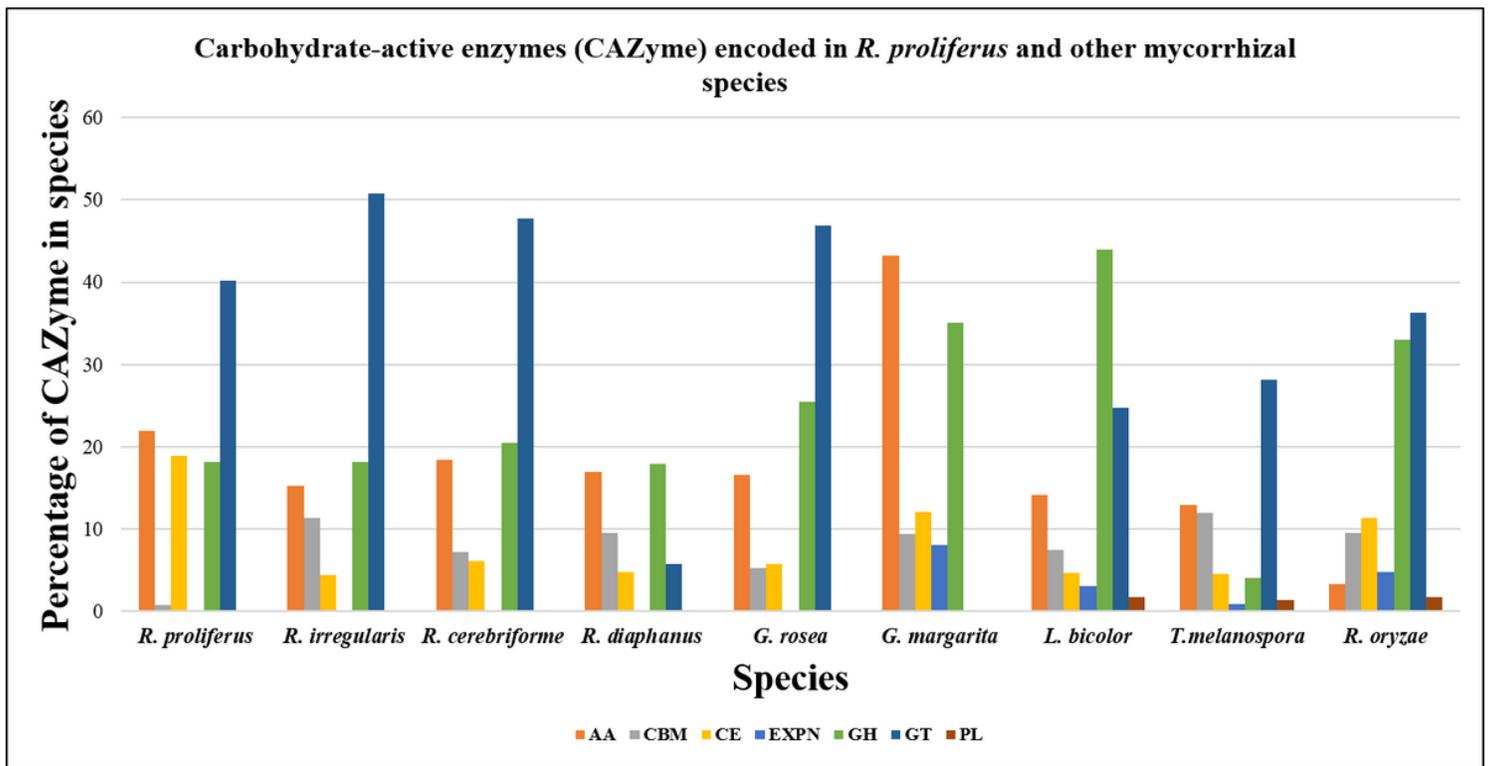


Figure 3

Distribution of important classes (AA, CBM, CE, GH, GT, PL and EXPN) under CAZyme in *R. proliferus* and other fungi

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.docx](#)
- [TableS4.docx](#)
- [TableS5.docx](#)
- [TableS6.docx](#)
- [TableS7.docx](#)
- [TableS8.docx](#)