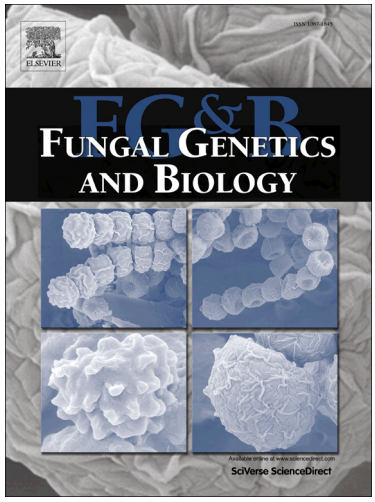


Sequencing and functional analysis of the genome of a nematode egg-parasitic fungus, *Pochonia chlamydosporia*

Eduardo Larriba, María D.L.A. Jaime, José Carbonell-Caballero, Ana Conesa, Joaquín Dopazo, Corey Nislow, José Martín-Nieto, Luis Vicente Lopez-Llorca

PII: S1087-1845(14)00018-8
DOI: http://dx.doi.org/10.1016/j.fgb.2014.02.002
Reference: YFGBI 2648

To appear in: *Fungal Genetics and Biology*

Received Date: 19 December 2013

# Sequencing and functional analysis of the genome of a nematode egg-parasitic fungus, *Pochonia chlamydosporia*

**Eduardo Larriba[1,2], María D.L.A. Jaime[3], José Carbonell-Caballero[4], Ana Conesa[4], Joaquín Dopazo[4], Corey Nislow[5,6,7]*, José Martín-Nieto[2,8] and Luis Vicente Lopez-Llorca[1,2§]**

[1]Department of Marine Sciences and Applied Biology, University of Alicante, P.O. Box. 99, E-03080 Alicante, Spain

[2]Multidisciplinary Institute for Environmental Studies (MIES) "Ramón Margalef", University of Alicante. Alicante, Spain.

[3]Department of Cell and Systems Biology, University of Toronto, M5S 3G5 Mississauga, Ontario, Canada.

[4]Institute for Computational Genomics, Príncipe Felipe Research Center, Avda. Eduardo Primo Yúfera 3, E-46012 Valencia, Spain.

[5]Department of Molecular Genetics, University of Toronto, M5S 1A8 Toronto, Ontario, Canada.

[6]Banting and Best Department of Medical Research, University of Toronto, Toronto, M5S 3E1 Toronto, Ontario, Canada.

[7]Terrence Donnelly Centre for Cellular and Biomedical Research, University of Toronto, M5S 3E1 Toronto, Ontario, Canada.

[8]Department of Physiology, Genetics and Microbiology, University of Alicante, E-03080 Alicante, Spain.

*Current address: Department of Pharmaceutical Sciences, University of British Columbia, V6T 1Z3 Vancouver, Canada

**[§]Corresponding author**

Luis Vicente Lopez-Llorca. Department of Marine Sciences and Applied Biology, University of Alicante, P.O. Box. 99, E-03080 Alicante, Spain. +34 965 90 3400 x 3381. lv.lopez@ua.es

Email addresses:

EL: eduardo.larriba@ua.es.

MJ: maria.jaime@gmail.com

JCC: jcarbonell@cipf.es

AC: aconesa@cipf.es

JD: jdopazo@cipf.es

CN: corey.nislow@gmail.com

JMN: jmnieto@ua.es

LVLL: lv.lopez@ua.es

**Abstract**

*Pochonia chlamydosporia* is a worldwide-distributed soil fungus with a great capacity to infect and destroy the eggs and kill females of plant-parasitic nematodes. Additionally, it has the ability to colonize endophytically roots of economically-important crop plants, thereby promoting their growth and eliciting plant defenses. This multitrophic behavior makes *P. chlamydosporia* a potentially useful tool for sustainable agriculture approaches. We sequenced and assembled ~41 Mb of *P. chlamydosporia* genomic DNA and predicted 12,122 gene models, of which many were homologous to genes of fungal pathogens of invertebrates and fungal plant pathogens. Predicted genes (65%) were functionally annotated according to Gene Ontology, and 16% of them found to share homology with genes in the Pathogen Host Interactions (PHI) database. The genome of this fungus is highly enriched in genes encoding hydrolytic enzymes, such as proteases, glycoside hydrolases and carbohydrate esterases. We used RNA-Seq technology in order to identify the genes expressed during endophytic behavior of *P. chlamydosporia* when colonizing barley roots. Functional annotation of these genes showed that hydrolytic enzymes and transporters are expressed during endophytism. This structural and functional analysis of the *P. chlamydosporia* genome provides a starting point for understanding the molecular mechanisms involved in the multitrophic lifestyle of this fungus. The genomic information provided here should also prove useful for enhancing the capabilities of this fungus as a biocontrol agent of plant-parasitic nematodes and as a plant growth-promoting organism.

**Keywords**:

*Pochonia chlamydosporia*, genome sequencing, nematophagous fungus, endophyte, hydrolytic enzymes, gene ontology.

**Highlights**

We present the first draft genome of a fungal nematode-egg parasitic fungus.

A close phylogenomic relationship exists between *P. chlamydosporia*, entomopathogenic and endophytic fungi.

The *P. chlamydosporia* genome is enriched in genes encoding hydrolytic enzymes.

Gene families in the *P. chlamydosporia* genome support its tritrophic lifestyle.

**Abbreviations**

PHI: Pathogen-host interaction; CAZy: Carbohydrate-active enzyme; MEROPS: Peptidase database; GH: Glycoside hydrolase; CE: Carbohydrate esterase; PKS: Polyketide synthase; NRPS: Non-ribosomal peptide synthetase; HYBRID: Hybrid PKS-NRPS enzyme; MFS: Major facilitator superfamily; ABC: ATP-binding cassette; CYP: Cytochrome P450; GPCR: G-protein coupled receptor; PK: Protein kinase; HK: Histidine kinase; TF: Transcription factor.

## 1. Introduction

Plant-parasitic nematodes cause an estimated global loss of US $100 billion annually (Casas-Flores and Herrera-Estrella, 2007). Fungi that parasitize and infect female nematodes and destroy their eggs include the worldwide-distributed egg-parasitic fungus *Pochonia chlamydosporia* (syn. *Verticillium chlamydosporium*, teleomorph *Metacordyceps chlamydosporia*) (Sung et al., 2007; Zare et al., 2007). *P. chlamydosporia* is a well known soil fungus (Domsch et al., 1993) which infects nematode females and eggs of economically important plant parasitic nematodes (such as *Heterodera* spp. and *Meloidogyne* spp.). In some agroecosystems this fungus is a main cause of soil suppressiveness to these nematodes acting as their agent of "natural" biological control. Since *P. chlamydosporia* is not an obligate parasite, it also acts as a true endophyte, colonizing plant roots of diverse species including main crops (such as barley and tomato). When switching host or habitats, this fungus faces a number of barriers and niches which requires diverse abilities (e.g. protein-glycan degradation), using hydrolases presumably encoded in its genome.

Infection of nematode eggs by *P. chlamydosporia* involves adhesion, differentiation of appressoria and egg-shell penetration (Lopez-Llorca et al., 2002b). We have recently documented the development of penetration hyphae in the *Meloidogyne javanica* egg-shell and invasion of egg contents by trophic hyphae using a GFP-tagged *P. chlamydosporia* strain (Escudero and Lopez-Llorca, 2012). However, information on the molecular basis of the infection process of nematode eggs by *Pochonia* species is scarce. There is evidence suggesting that secreted hydrolytic enzymes including proteases (Huang et al., 2004; Lopez-Llorca and Robertson 1992; Morton et al., 2003) and chitinases (Tikhonov et al., 2002) are fundamental for the degradation of egg-shell components. Rosso et al. (2011) used cDNA-amplified fragment length polymorphism (cDNA-AFLP) based on transcript profiling in order to identify genes involved in the pathogenesis of nematode eggs by *P. chlamydosporia*, and found that genes encoding transcription factors, transporters and enzymes involved in fungal metabolism are

enriched. In particular, expression of the *P. chlamydosporia* VCP1 serine protease is induced in response to nitrogen and carbon sources, environmental conditions and the presence of nematode eggs (Ward et al., 2012).

Similarly to other Hypocreales (such as *Trichoderma* spp. and *Beauveria bassiana*), *P. chlamydosporia* has a broad host range and can act as an endophyte in both monocot (Lopez-Llorca et al., 2002a) and dicot (Bordallo et al., 2002) plants. Interestingly, this fungus promotes plant growth in barley (*Hordeum vulgare*) (Macia-Vicente et al., 2009b), wheat (*Triticum aestivum*) (Monfort et al., 2005), tomato (*Solanum lycopersicum*) (Escudero and Lopez-Llorca, 2012), lettuce (*Lactuca sativa*) (Dias-Arieira et al., 2011) and pistachio (*Pistacia vera*) (Ebadi et al., 2009). *P. chlamydosporia* colonizes the rhizoplane of crop plants, especially cereals, forming abundant chlamydospores (Bordallo et al., 2002; Kerry, 2000). This fungus then penetrates root hairs and epidermal cells, and colonizes the cortex but not the root vascular system (Bordallo et al., 2002; Macia-Vicente et al., 2009a). *P. chlamydosporia* modulates biochemical (e.g. biosynthesis of secreted phenolic compounds) and structural (e.g. root papillae formation) plant defenses (Bordallo et al., 2002; Escudero and Lopez-Llorca, 2012; Macia-Vicente et al., 2009a), but these do not prevent root colonization by the fungus. *P. chlamydosporia* expresses proteases during this process, such as VCP1 and SCP1, the latter a newly reported serine carboxypeptidase (Larriba et al., 2012; Lopez-Llorca et al., 2010).

*P. chlamydosporia* also produces several polyketide compounds named pochonins (Hellwig et al.,2003), which are resorcylic acid lactones derived from radicicol. Different multifunctional polyketide synthases are responsible for carbon skeleton construction, acting at the initial steps in the pochonin biosynthesis pathway (Reeves et al., 2008). Despite the potential of pochonins in biotechnology and pharmacology (Barluenga et al., 2009), only two genes encoding polyketide synthases from *P. chlamydosporia* have been cloned and characterized so far (Reeves et al., 2008; Zhou et al., 2010).

The sequencing and functional analysis of genomes recently available from fungal pathogens of invertebrates (Gao et al., 2011; Zheng et al., 2011) have increased our knowledge on the biocontrol capabilities of these organisms. The in-depth functional genomic characterization of the nematophagous fungus *P. chlamydosporia* presented in this work provides new insights into the molecular mechanisms enabling its multitrophic lifestyle as a saprophyte, nematode pathogen and endophyte.

## 2. Materials and methods

### 2.1. Fungal strains

*Pochonia chlamydosporia* isolate 123 (ATCC number MYA-4875) grown on corn meal agar (CMA) at 25 ºC in the dark for 1–4 weeks was inoculated in four flasks each containing 50 ml of potato dextrose broth (PDB) and incubated at 25º C for 8 days with shaking at 120 rpm.

### 2.2. Fungal genome sequencing and assembly

Isolation of genomic DNA was carried out using the DNeasy Plant Mini Kit (Qiagen) according to the manufacturer´s fungal protocol. Total DNA obtained was subjected to quality control by agarose gel electrophoresis and quantified by the same method. The genome of *P. chlamydosporia* was sequenced with MPS (massively parallel sequencing) Illumina technology. Two DNA libraries were constructed: a pair-end library with an insert size of 200 bp and a mate-pair library with an insert size of 3 kb. Each DNA library was sequenced using an Illumina HiSeq 2000 at the Donnelly Sequencing Centre (University of Toronto). Pair-end and mate-pair sequences were analyzed using the FastQC application (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc) for quality control. In order to improve the quality of the sequences, Illumina PCR adapter sequences were removed using the Cutadap tool (Martin, 2011), and sequences were then futher trimmed based on the Phred quality score (>20) using DynamicTrim (Cox et al., 2010). The genomic trimmed sequences were assembled by SOAPdenovo (http://soap.genomics.org.cn/soapdenovo.html) using mate-pair reads to generate scaffolds. Transfer RNA (tRNA) genes were predicted with tRNAscan-SE (Lowe and Eddy 1997), and genome repetitive sequences were analyzed using RepeatMasker (http://www.repeatmasker.org/). Repetitive elements were identified by RMBlastN search against RepBase v. 20120418 (http://www.girinst.org/repbase). This Whole Genome Shotgun project has been deposited at the DDBJ/EMBL/GenBank database under the accession number AOSW00000000.

2.3. Gene prediction, annotation and protein classification

*Ab initi*o gene prediction was performed on the *P. chlamydosporia* genome assembly by using the Augustus program (Stanke et al., 2006), and as evidence genes of this fungus available in the GenBank database (consulted in June 2013). ESTs from *P. chlamydosporia* (Rosso et al., 2011) and *M. anisopliae* (Wang et al., 2005; Zhang and Xia 2008, 2009), and the training annotation files from *F. graminearum* were incorporated as references. The GeneMark-ES program (Ter-Hovhannisyan et al., 2008) was used as a self-training *ab initio* predictor. A final set of gene models was selected by MAKER2 (Holt and Yandell, 2011) by combining *ab initi*o predictions with transcripts listed above aligned using Exonerate (http://www.ebi.ac.uk/~guy/exonerate). Consensus sets of gene models were functionally annotated using Blast2GO (http://www.blast2go.com/b2ghome). Gene families were established using the Interpro database (http://www.ebi.ac.uk/interpro), and putatively-secreted proteins were identified using SignaIP (http://www.cbs.dtu.dk/services/SignalP), both implemented in the Blast2GO suite. Proteases were classified into families by means of batch Blast search against the MEROPS database (Rawlings and Morton, 2008). Carbohydrate-active enzymes (CAZymes) were classified by performing a HMMER (http://eddylab.org) scan based on the profiles compiled in dbCAN release 2.0 on the CAZy database (http://www.cazy.org). To identify potential pathogenesis genes, a whole genome Blast search (E-value $<10^{-6}$) was conduced against the Pathogen Host Interactions (PHI) database v. 3.2 (http://www.phi-base.org). Secondary metabolism genes were predicted using SMURF (http://jcvi.org/smurf/index.php).

2.4. Transcriptomic analysis by RNA-Seq

Barley roots were inoculated with *P. chlamydosporia* as described in Lopez-Llorca et al. (2010). Isolation of RNA from barley roots inoculated with the fungus and non-colonized barley roots was carried out using the RNeasy Plant Mini Kit (Qiagen). Subsequently, the RNA obtained was amplified using the MessageAmp II kit (Ambion, Life Technologies) according to the

manufacturer's instructions. Amplified RNA obtained was reverse-transcribed into cDNA and sequenced using a Illumina Genome Analyzer at Macrogen Inc. (South Korea). Pair-end sequencing yielded 77,606,102 reads from control roots and 99,966,600 reads from inoculated roots, all of them with an average length of 101 bp. All reads were aligned against the genome of *P. chlamydosporia* using Bowtie (Langmead et al., 2009), and 544,060 reads (0.54% of total) from inoculated plants RNA were found to map with the genome of *P. chlamydosporia*. On the contrary, zero reads from control roots RNA mapped against the fungus genome. To validate the gene models predicted in the *P. chlamydosporia* genome, the genes expressed during fungal endophytic colonization of barley roots were identified by employing a combination of TopHat (Trapnell et al., 2009) and HTSeq v0.5.4p2 (http://www-huber.embl.de/users/anders/HTSeq) software. These genes were annotated and classified as described in section 2.3.

2.5. Orthology and phylogenomic analysis

Orthologous gene detection in *P. chlamydosporia* and 28 additional fungal genomes was carried out using ProteinOrtho (Lechner et al., 2011). In total, 382 orthologous sequences were identified, acquired and aligned using T-Coffee (Notredame et al., 2000). Alignments of orthologous sequences were concatenated using FASconCAT (Kück and Meusemann, 2010), and a maximum likelihood phylogenomic tree was created with the program TREE-PUZZLE using the Dayhoff model (Schmidt et al., 2002). The tree obtained was edited using the TreeGraph 2 software (Stöver and Müller, 2010).

## 3. Results

3.1. Genome sequencing and general features

*P. chlamydosporia* strain 123 (ATCC MYA-4875) was found suppressing a cereal cyst nematode (*Heterodera avenae*) population in a field near the city of Seville (SW Spain). The genome of this fungus was sequenced 136-fold coverage using a whole-genome shotgun approach and Illumina sequencing technology. The completed assembly (N50, 225 kb) comprised 901

scaffolds and 57 additional contigs >3 kb (not assembled into scaffolds) which, after discarding ambiguous bases, yielded a final genome size of 41.2 Mb. The main features of this fungal genome are summarized in Table 1. From this assembly we predicted a total of 12,122 gene models using MAKER2 (see section 2.3), a coding capacity that was similar to that of other ascomycetes (Table 2). A 63% fraction of these gene models was expressed, i.e. validated using RNA-Seq, under endophytism (Supplementary Table 1). A 5.5% (672) of total predicted genes in the genome did not exhibit any homologous counterparts in the NCBI database (accessed June 2013), of which ca. 41% (277 genes) were expressed during *P. chlamydosporia* endophytic lifestyle. Protein-coding genes of this fungus showed the highest number of homologous proteins (34% and 30%, respectively) with the entomopathogenic fungi *M. anisopliae* and *M. acridum* (Table 2).

We built a phylogenomic tree based on *P. chlamydosporia* genome-encoded orthologous proteins found in the genomes of other 28 filamentous fungi and yeasts (Fig. 1). This tree illustrated that *P. chlamydosporia* was most closely related to entomopathogenic fungi *M. anisopliae* and *M. acridum*. The endophyte *Epichloë festucae*, also belonging to the Clavicipitaceae, formed a clade with both the nematode egg-parasite and the two entomopathogens. Other relevant entomopathogenic fungi (*B. bassiana* and *Cordyceps militaris*) formed an independent, but closely related clade, whereas mycoparasitic *Trichoderma* spp. and plant-pathogenic *Fusarium* spp. form adjacent, independent clades (Fig. 1).

3.2. Mobile elements

A 0.46% fraction of the *P. chlamydosporia* genome consisted of repeated sequences, using the RepBase database, we were able to identify 209 retrotransposons (class I), 133 DNA transposons (class II), and 7 unknown elements (Fig. 2A). The number and family distribution of mobile elements of *P. chlamydosporia* compared to those of insect pathogenic fungi are shown in Fig. 2B. This comparison showed that the *P. chlamydosporia* genome exhibited more copies of most

of the 7 families of mobile elements than did the closely-related insect pathogenic fungi *Metarhizium* spp.

### 3.3. Gene Ontology analysis of protein-coding genes

We functionally annotated 7,887 protein-coding genes, representing 65% of the *P. chlamydosporia* whole set of genes in the genome. The distribution of these genes into functional groups according to Gene Ontology (GO) is shown in Fig. 3. Within the Biological Process GO domain (Fig. 3A), protein-coding genes distributed into the following GO terms: ca. 21% were associated with primary metabolism, 11% with macromolecule metabolism, such as protein synthesis and modification, 6% of annotated genes with catabolism, 9% with biosynthetic processes and 7% with nitrogen metabolism (Supplementary Fig. 1A). Within the metabolism GO term, 13% of the annotated genes belonged to the cellular metabolic processes GO term (Fig. 3A), whereas 17% of genes were related to other cellular processes, such as cell cycle regulation, cell growth, cell communication and cell death (Supplementary Fig. 1B). Finally, 2% of genes were related to stress responses, and 1% were annotated within interspecies interactions between organisms (Fig. 3A). Within the Molecular Function GO domain (Fig. 3A), we found that 23% of annotated genes encoded proteins related to hydrolase activity, breaking of phosphorus ester bonds, acid anhydrides and peptides, while a further 21% of genes were related to transferase activity, including protein kinases (Supplementary Fig. 1C). A larger set of genes (43%) were associated with binding of molecules and ions (Supplementary Fig. 1D), with 23% of them being involved in binding to small molecules and 16% in nucleic-acid binding, and transcription factors representing an additional 2% of annotated genes (Fig. 3A).

### 3.4. Pathogenesis-related genes

To identify genes involved in pathogenicity in the *P. chlamydosporia* genome, we conducted a Blast search against the Pathogen-Host Interaction (PHI) database, which compiles experimentally-validated pathogenesis-related genes of fungi, bacteria and oomycetes

(Winnenburg et al., 2008). We found that 16% of the protein-coding genes in the genome of *P. chlamydosporia* (1,981 genes) shared homology with genes present in the PHI database (Supplementary Fig. 2A), 24% of which (468 genes) coded for putatively secreted polypeptides. Despite the absence of nematophagous fungi in this database, we considered experimental validation of pathogenesis for a gene in a given fungal species as suggestive of a pathogenic role for its homolog in other fungi. Using this criterion, virulence genes have been analyzed in the entomopathogenic fungi *M. anisopliae* and *M. acridum* (Gao et al., 2011). In order to identify the biological processes and molecular functions attributable to the virulence genes identified using the PHI database, we carried out a functional annotation using Blast2GO (Fig. 3B). Genes identified within the Biological Process GO domain fell into 28 functional subcategories, which included genes related to primary metabolism (20%), establishment of localization (15%) and cellular processes (13%). Likewise, regulation of biological processes (7%), biosynthetic processes (6%) and catabolism (7%) were represented in that domain. Interestingly, of the 81 protein-coding genes identified as associated with the GO term relationships between organisms, i.e. pathogenesis/symbiosis (Fig. 3A), 47 of them were found in the PHI database (Fig. 3B and Supplementary Fig. 2A). This set of genes included those coding for glycoside hydrolases, proteases, proteins involved in signal transduction, detoxification processes and stress response factors (Supplementary Fig. 2B, 2C and Supplementary Table 2). Within the Molecular Function GO domain (Fig. 3B), we found that nucleotide-binding was the most represented GO term (31%), and a large proportion of the annotated genes to encode proteins having hydrolase activity (14%), transferase activity (16%) or nucleic acid-binding properties (12%), with an additional 3% being related to transcription factors (Fig. 3B).

### 3.5. Transcriptomic analysis of root endophytic colonization

In order to identify protein-coding genes involved in *P. chlamydosporia* endophytic behavior, an RNA-Seq analysis was carried out using fungus-colonized barley roots. Using this technique we identified a set of 7,586 genes expressed from the *P. chlamydosporia* genome (Supplementary

Table 1). Fifty-seven percent of all *P. chlamydosporia* genes coding for predicted secreted proteins (1,432) were detected to be expressed during endophytism. Within the Biological Process GO domain the genes expressed by this fungus were associated with a total of 28 GO terms, including primary metabolic process (21% of the genes), cellular metabolic process (19%) and nitrogen compound metabolic process (Fig. 3C). Likewise, regulation of biological processes (6%), biosynthetic processes (9%) and macromolecule metabolic process (6%) were represented GO terms. Within the Molecular Function GO domain we found that the most represented terms were related to small molecule binding and hydrolase activity. Likewise, transferase activity, binding to nucleic acids, and protein binding were each represented by a 16% fraction of expressed genes (Fig. 3C).

The genes found to be expressed in our transcriptomic analysis of barley root colonization by *P. chlamydosporia* were compared with those exhibiting a homolog within the PHI database (putatively associated with pathogenesis) and with those annotated under the interaction between organisms GO term, and a common set of 32 genes was obtained (Supplementary Fig. 2A). The set of genes putatively associated with pathogenesis plus endophytism is shown in Supplementary Fig. 2B and Supplementary Table 2, the majority of which encoded hydrolytic enzymes and signal transduction proteins. Hydrolases found to have homologs in the PHI database mostly included metalloproteases and a chitinase precursor, whereas those expressed under endophytism were more diverse, including serine and rhomboid protease families and a protein phosphatase (Supplementary Fig. 2D and supplementary Table 2).

3.6. Hydrolytic enzymes encoded by the *P. chlamydosporia* genome

The nematode egg-shell constitutes the main barrier against parasitism by nematophagous fungi. *P. chlamydosporia* employs an array of hydrolytic enzymes, including proteases, chitinases, esterases and lipases, in order to degrade and penetrate nematode egg-shells (Huang et al., 2004; Yang et al., 2007). As illustrated in Fig. 4, its genome contained an ample set of genes putatively

encoding hydrolytic enzymes, over half of which were detected to be expressed during root endophytic colonization (Supplementary Table 3).

3.6.1. Proteases

We carried out a batch Blast search against the complete MEROPS protease database (Rawlings and Morton, 2008) and found 522 genes coding for proteases, which were classified into six categories according to their catalytic type, and distributed in a total of 68 families (Supplementary Table 3). Serine proteases (189 genes) were the largest category of proteases encoded in the *P. chlamydosporia* genome (Fig. 4), with 59% of genes in this family being expressed during endophytism (Fig. 4). Serine proteases constituted the second largest family of hydrolytic enzymes expressed under endophytism (Fig. 5A). Among these enzymes, subtilisins and serine carboxypeptidases are known to be involved in nematode egg-parasitism (Lopez-Llorca et al., 2002b; Ward et al., 2011) and expressed during root endophytic colonization (Lopez-Llorca et al., 2010) by this fungus. We found in this context that the *P. chlamydosporia* genome coded for 32 serine proteases of the S8 family (subtilisins) and 16 proteases of the S10 family (serine carboxypeptidases), 7 of the latter being putatively secreted enzymes (Fig. 5B). Fifty percent of genes in the S10 family were expressed during root endophytism (Figure 5B and Supplementary Table 3). A Blast search against the PHI database returned a number of homologous genes coding for putatively secreted serine proteases belonging to other fungi (Supplementary Table 4). Among these genes, the *P. chlamydosporia* genome encodes two S54 (rhomboid family) membrane proteases (Fig. 5B), identified in this study as putatively involved in endophytic capacity (Supplementary Tables 2 and 3). Also during endophytism, we detected the expression of 55% of the S33 family members, this being the family of serine proteases (prolyl aminopeptidases) with the highest number of members identified in the genome (Supplementary Table 3). The second largest group of proteases in this genome was metalloproteases, with 147 genes belonging to 25 families (Supplementary Table 3), of which

71% were identified in our transcriptomic analysis (Fig. 4). The largest family within metalloproteases was glutamate carboxypeptidases, M20 (Fig. 5B). The *P. chlamydosporia* genome contained genes encoding 3 enzymes of the metalloproteases M36 (fungalysins) family, 9 of the M28 (aminopeptidase Y-related), 6 of the M35 (deuterolysins) and 14 of the M43 (cytophagalysins) (Fig. 5B), with representatives of all of them being expressed during endophytism (Fig. 5B and Supplementary Table 3). Other abundant families of putatively secreted proteases were aspartic (A1) and cysteine proteases (Fig. 4 and 5B), of which we detected expression of many of its members during root endophytism (Fig. 5A and Supplementary Table 3).

3.6.2. Glycoside hydrolases

Besides proteins, *P. chlamydosporia* needs to degrade structural polysaccharides of nematode egg-shells, such as chitin. Using CAZy database models (see section 2.3) we identified 305 genes encoding glycoside hydrolases (GH) and 159 encoding carbohydrate esterases (CE) in its genome, 200 and 54 of which, respectively, were putatively secreted (Fig. 4A and Supplementary Table 3). During endophytic root colonization 60% and 39% of the genes identified as GH and CE, respectively, were expressed (Fig. 5B and Supplementary Table 3). Among hydrolases the GH group included the largest number of genes expressed during endophytism (Fig. 5A). Chitinases (GH18) were the most abundant *P. chlamydosporia* glycoside hydrolases, represented by 22 genes, of which 15 were putatively secreted (Fig. 5B and Supplementary Table 3). These enzymes degrade the chitin present in the chitin–protein complex of the nematode egg-shell (Bird and McClure, 1976). Regarding chitosanases (GH75), the genome of this fungus encoded 11 enzymes of this family, 10 of which were putatively secreted (Fig. 5A). Of the enzymes related with chitin/chitosan degradation, i.e. GH18 plus GH75 families, less than 40% of their members were expressed during root endophytism (Fig. 5B). GH76 (α1,6-mannanases) constituted the family exhibiting the highest number of putatively

secreted hydrolases, with 14 encoding genes (Fig. 5B). Genes encoding members of cellulase families (GH5-GH12) were also expressed during root colonization (Fig. 5B) and exhibited homology with enzymes in the PHI database (Supplementary Table 4). We were also able to identify expressed GH families involved in the degradation of callose, lignocellulose and xylans (Fig. 5B), or in changes in cell wall composition and morphogenesis (GH17, GH31, GH72 and GH125) in fungal pathogens ( Fig. 5B, Supplementary Tables 2 and 3).

3.6.3. Carbohydrate esterases

The largest family CE genes contained in the genome of *P. chlamydosporia* was sterol esterases (CE10), represented by 93 genes, 29 of which were predicted to encode putatively secreted enzymes and 36% of which were expressed during root endophytic colonization (Fig. 5B and Supplementary Table 3). Its genome encoded 9 cutinases (CE5 family) with homologous counterparts in the PHI database (Fig. 5B and Supplementary Table 4). We also detected 27 genes encoding triglyceride lipases and 11 phospholipases (Fig. 5B) with homologs in the PHI database (Supplementary Table 4). Only 18 % phospholipase genes were expressed during endophytic behavior, in contrast to the high number of triglyceride lipase genes (82%) expressed during this phase (Fig. 5A).

3.7. Secondary metabolism gene clusters

*P. chlamydosporia* produces a variety of secondary metabolites, such as radicicol (=monorden), tetrahydromonorden, pseurotin A, pochonins A to J (Hellwig et al., 2003, Shinonaga et al., 2009 and Zhou et al., 2010) and various aurovertin-type metabolites (Niu et al., 2010). In this context, we found that its genome contained genes putatively encoding 15 polyketide synthases (PKS) and 12 putative non-ribosomal peptide synthases (NRPS), together with a number of PKS and NRPS-like proteins and 4 NRPS-PKS hybrid genes (Supplementary Table 5). We also found a radicicol gene cluster, *Rdc1-Rdc5* (Zhou et al., 2010) (Supplementary Table 6). During

endophytic root colonization, *P. chlamydosporia* expressed 56% of the secondary metabolism pathway core genes identified in its genome (Supplementary Table 5). Among them we detected the expression of seven genes related to the radicicol cluster (Supplementary Table 6).

3.8. Genes involved in transport, detoxification and cell wall modification

The *P. chlamydosporia* genome encoded 290 transporters of the major facilitator superfamily (MFS), 58 ATP-binding cassette (ABC) transporters and 113 general transporters (Fig. 4 and Supplementary Table 7), most of which exhibited homologs in the PHI database (Supplementary Table 5). During endophytic root colonization *P. chlamydosporia* expressed genes encoding drug resistance, sugar/inositol, oligopeptide and amino acid transporters (Fig. 4 and Supplementary Table 8). Also, we detected several genes encoding oxidoreductases, many of them related to detoxification (Fig. 4 and Supplementary Table 8). These comprised 110 cytochrome P450 (CYP) enzymes, a number similar to those found in the entomopathogen *M. anisopliae* and in fungal plant pathogens (Supplementary Table 7 ). Interestingly, the number of members of these oxidoreductase families, which are likely involved in counteracting oxidative stress, was higher in *P. chlamydosporia* than in most invertebrate and plant pathogens (Supplementary Table 7). Sixty percent of CYP-coding genes in the *P. chlamydosporia* genome were expressed during endophytism (Fig. 4) and 72% showed homology with genes in the PHI database (Supplementary Table 8). Finally, the *P. chlamydosporia* genome included genes coding for enzymes involved in cell wall biosynthesis and modification, such as chitin synthesis activators (Sel-1 domain-containing proteins), chitin synthases, lipopolysaccharide-modifying proteins and hydrophobins (Fig. 4), which were expressed during endophytic colonization (Fig. 4 and Supplementary Table 8).

3.9. Signal transduction and regulation of gene expression

In order to adjust to its different lifestyles (saprophytic, parasitic and endophytic), *P. chlamydosporia* needs genes involved in signaling and regulation of gene expression. G proteins are involved in different biological processes in filamentous fungi, including development and pathogenesis, and are responsible as well of transducing environmental signals. The genome of this fungus encoded eight G-protein subunits (Fig. 4), three of which showed homology with G proteins involved in vegetative growth, conidiation, conidium attachment, appressorium formation, mating, and pathogenicity in *Ma. oryzae* (Liu and Dean, 1997). Six of these G protein-coding genes were expressed during *P. chlamydosporia* endophytism (Fig. 4 and Supplementary Table 8). The genome of this fungus encoded 54 proteins homologous to Pth11-like G protein-coupled receptors (GPCR) of *Magnaporthe* spp., i.e. the same number found in the *M. anisopliae* genome (Supplementary Table 7). The *P. chlamydosporia* genome encoded as well 27 small GTPase regulators, (96% of which had homology with genes in the PHI database) and 12 Rab GTPase activators, all of which presumably modulate its endophytic behavior (Fig. 4 and Supplementary Tables 2 and 8). Among proteins involved in the regulation of cell and metabolic processes, the *P. chlamydosporia* genome encoded 153 protein kinases (PKs), 75% of which returned homologous partners within the PHI database (Fig. 4 and Supplementary Table 8). This number was similar to that found in entomopathogens and plant pathogens, except for *Ma. oryzae* (Supplementary Table 7). Together with PKs, we found genes coding for histidine kinases (HKs) in the genome of *P. chlamydosporia* (Fig. 4), nearly all which had homologous partners in the PHI database and 14 of which were expressed in endophytism (Fig. 4 and Supplementary Table 8). Finally, its genome encoded 409 putative transcription factors (TFs) grouped into six families (Fig. 4), of which that containing the highest number of genes expressed under endophytism was the $Zn_2Cys_6$ fungal type TF family (Fig. 4 and Supplementary Table 8).

## 4. Discussion

In this study we used MPS techniques in combination with bioinformatic tools (Jackman and Birol, 2010; Yandell and Ence, 2012) to obtain the first genome sequence of a nematode-egg parasitic fungus. In addition, we used RNA-Seq to validate our bioinformatically-predicted gene models, this being the first transcriptomic study of root endophytic colonization by a nematophagous fungus. *P. chlamydosporia* uses appressoria for host penetration (Escudero and Lopez-Llorca, 2012; Lopez-Llorca et al., 2002b), just like entomopathogenic (St. Leger et al., 1991) and plant-pathogenic fungi (Tucker and Talbot, 2001). On the contrary, nematode-trapping fungi, such as *Arthrobotrys oligospora*, generate complex hyphal networks and constrictive rings to capture motile nematodes (Liu et al., 2012; Yang et al., 2011). These differences in pathogenesis mechanisms are reflected in the higher homology we have found between *P. chlamydosporia* predicted protein-coding genes and those of entomopathogenic fungi (*Metarhizium* spp.) compared to those of the nematode-trapping fungus *A. oligospora*. This would support the idea that the genomic machineries of nematode egg-parasites and trapping fungi are very different, and that the nematophagous habit evolved independently several times in the Fungi. Our phylogenomic tree shows the existence of a close relationship between *P. chlamydosporia*, entomopathogenic, mycoparasitic (*Trichoderma* spp.) and -especially- endophytic fungi (*Epichloë festucae*). However, other specialized plant pathogens (*Verticillium* spp. wilt fungi) are less related to *P. chlamydosporia*. This confirms that the former *Verticillium* genus, which included nematode-egg (such as *P. chlamydosporia*) and insect parasites, plant pathogens and soil saprophytes, was artificially based on morphological features (mainly conidiophore morphology). Taken together, these results support the taxonomic affiliation of *P. chlamydosporia* as well as its ecology as endophyte and invertebrate pathogen (nematode-egg parasite). Also, they provide evidence of strong links with entomopathogenic fungi and of an existing evolutionary distance with nematode-trapping fungi. The present genome-wide study is also consistent with our recent sequence comparative analysis of VCP1 and P32 serine proteases

from the closely related egg-parasites *P. chlamydosporia* and *P. rubescens*, respectively, which exhibited higher similarity to proteases of fungal pathogens of insects than to those of nematode-trapping fungi (Larriba et al., 2012). The multitrophic lifestyle of *P. chlamydosporia*, i.e. as soil saprophyte, nematode-egg pathogen and root endophyte, is reflected in the large number of putatively secreted enzymes encoded by its genome, being the species showing the highest number of these among all ascomycete genomes sequenced to date (Galan et al., 2005: Gao et al., 2011; Islam et al., 2012; Xiao et al., 2012). More than half of them are expressed during *P. chlamydosporia* endophytic behavior. In this context, fungal adaptation to several lifestyles is thought to require a large number of hydrolytic enzymes and transporters for an efficient use of the diverse nutrients available (Gao et al., 2011). Accordingly, 30% of the *P. chlamydosporia* genes annotated within the Molecular Function GO domain are involved in the hydrolysis of a wide array of substrates, which is one of the most represented gene sets expressed during root endophytism. For instance, *P. chlamydosporia* genome encodes more proteases (and putatively secreted proteases) than the fungal insect pathogens *M. anisopliae* and *M. acridum* (Gao et al., 2011). The most studied proteases in *P. chlamydosporia* are those belonging to the serine protease category, among which subtilisins (S8) are the family with the highest number of putatively secreted members in its genome. Within this family, the VCP1 protease (Morton et al., 2003) has been found to be involved in degradation and penetration of the nematode egg-shell (Lopez-Llorca et al., 2002b) and to participate in the endophytic phase of the fungus (Lopez-Llorca et al., 2010). It must be mentioned that serine proteases are well known pathogenic determinants of both entomopathogenic (St. Leger, 1995) and human pathogenic fungi (Monod, 2008). Prolyl aminopeptidases (S33) are the most abundant serine protease family in the *P. chlamydosporia* genome. To this respect, proline residues constitute a large fraction (35%) of the amino acids of proteins forming the Tylenchida family egg-shell (to which plant-parasitic nematodes such as *M. javanica* belong) (Bird and McClure 1976). Likewise, the finding of *P. chlamydosporia* homologous protease genes in the PHI database related to the evasion of host

defenses, suggests that its proteases may play roles other than nematode egg-shell disruption (Hung et al., 2005; Jia et al., 2000; Newport et al., 2003; Soloviev et al., 2011). To this respect, the nematode *Meloidogyne incognita* genome contains genes related to the immune response and antifungal defenses (Abad et al., 2008). Expression of the VCP1 protease and the serine carboxypeptidase SCP1 has been detected during *P. chlamydosporia* endophytism (Lopez-Llorca et al., 2010). Additionally, we have identified in this work several families of expressed proteases putatively related to fungal-plant interaction, such as the rhomboid protease family (S54), likely involved in its endophytic behavior. Alongside, metalloproteases M28, M35, M36 and M43 expressed during *P. chlamydosporia* endophytism have been related to root endophytism by the basidiomycete *Piriformospora indica*, presumably by allowing degradation of plant cell wall proteins (Zuccaro et al., 2011).

Chitin is a major component of nematode egg-shells (Bird and McClure, 1976; Warton, 1980), and *P. chlamydosporia* is known to express chitinases for egg penetration (Gortari and Hours 2008; Mi et al., 2010; Tikhonov et al., 2002). Our *P. chlamydosporia* genome analysis has shown that chitinases (GH18) are the most represented family of GH, the majority of them predicted to be putatively secreted. Chitinases and chitosanases are coexpressed with protease VCP1 by *P. chlamydosporia* in the presence of chitin (Palma-Guerrero et al., 2008), perhaps suggesting their co-involvement in the pathogenesis of nematode eggs. The number of GH enzymes encoded by a fungus genome has been linked to its capacity to adaptation to diverse environments (Van den Brink and De Vries, 2011). *P. chlamydosporia* expresses 60% of its GH encoding genes during endophytism. These include GH enzymes that have been related to the degradation of cellulose, hemicellulose, xylans and other constituents of the plant cell wall (Gibson, 2012). Additionally, this fungus exhibits a large number of CE enzymes, which would cleave ester bonds present in plant polysaccharides (Biely, 2012). Transporters that have been involved in the trophic behavior of fungal endophytes (Zuccaro et al., 2011) and during pathogenesis of nematode eggs (Rosso et al., 2011) are encoded in the *P. chlamydosporia* genome and exhibit homologs in the PHI

database, some of which are also expressed during endophytism. *P. chlamydosporia* also expresses a large number of genes related to sugar/inositol transport, which are involved in the establishment of plant-fungus relationships by *M. anisopliae* (Fang an d St. Leger 2010).

The multiple modes of life of *P. chlamydosporia* would require an extensive signaling machinery to perceive and respond to a wide variety of environmental stimuli. In this light, its genome contains more genes encoding G-protein αsubunits than the entomophatogen *M. anisopliae*, which are involved in the *P. chlamydosporia* endophytic phase, as well as a larger number of Pth11-like G-proteins involved in physiological processes in the plant pathogen *Ma. grisea* (DeZwaan et al., 1999). Several genes encoding protein kinases were also identified here in the *P. chlamydosporia* genome, a type of proteins that have been involved in the regulation of virulence genes in the entomopathogenic fungus *M. anisopliae* (Fang et al., 2009). Our analyses show that most families of genes coding for proteins related to signal sensing and transduction in *P. chlamydosporia* exhibit homologous genes in the PHI database and are expressed in endophytism. During the pathogenesis of nematode eggs, *P. chlamydosporia* is known to express transcription factors (Rosso et al., 2011) belonging to $Zn_2Cys_6$ fungal-type, bZIP and bromodomain-containing TF families, which are well represented in the *P. chlamydosporia* genome. These observations highlight the importance of identifying the signaling and transduction pathways involved in the transitions between diverse *P. chlamydosporia* lifestyles.

Nematode-egg parasitism and root endophytism by *P. chlamydosporia* requires a broad set of genes involved in detoxification and resistance to oxidative stress, such as cytochrome P450 genes, present in a similar number than in the *M. anisopliae* genome (Gao et al., 2011). The *P. chlamydosporia* genome encodes more monoxygenases that any insect-pathogenic fungal genome sequenced to date (Xiao et al., 2012), and which are presumably involved in its endophytic phase. Monoxygenases have been found in fungal oxidation of plant phenolic compounds, thereby reducing plant defenses (Morrissey and Osbourn, 1999). Besides, *P.*

*chlamydosporia* encodes a large number of zinc type alcohol dehydrogenases, wich are necessary for mannitol synthesis, a sugar associated with stress tolerance and energy storage in fungi (Solomon et al., 2007). Finally, the *P. chlamydosporia* genome includes genes for the synthesis of a large number of secondary metabolites (Hellwig et al., 2003; Niu et al., 2010; Shinonaga et al., 2009). Our analysis allowed us to identify the elements of a gene cluster involved in the biosynthesis of radicicol, a compound which is the structural basis for pochonin production (Reeves et al., 2008). Expression of PKS genes (necessary for polyketide biosynthesis) has been detected during the *P. chlamydosporia* endophytic and saprophytic phases (Rosso et al., 2011). There is also evidence that the production of secondary metabolites is involved in the pathogenesis of nematode eggs (Niu et al., 2010) and in adaptation to the root environment (Johnson et al., 2007).

**5. Conclusions**

In this work we show evidence that the genome of the nematode-egg parasite *P. chlamydosporia* is most closely related to those of Clavicipitaceous enthomopathogenic fungi (*Metarhizium* spp.). Both fungi form a clade with *Epichloë festucae*, an endophyte. This supports the presence of endophytism in both nematophagous and entomopathogenic fungi. In addition, the *P. chlamydosporia* genome is evolutionarily close to that of mycoparasitic (*Trichoderma* spp.) and plant pathogenic (*Fusarium* spp.) fungi. However, the nematode-trapping fungus *A. oligospora* is phylogenomically far apart from *P. chlamydosporia*. This would agree with the idea of independent and multiple evolution of nematophagous behavior in fungi. The wide array of hydrolytic enzymes and transporters encoded by the *P. chlamydosporia* genome support the observations of multitrophic behavior (pathogenic, endophytic and saprophytic) by *P. chlamydosporia* (Fig. 6). To this respect, this fungus expresses ca. 62% of its protein-coding genes during barley root endophytism. The number and diversity of genes related to signal transduction and gene regulation in its genome suggests that the transitions between the different trophic modes require a sophisticated signaling network (Fig. 6). This study provides the

"essential parts" list to understand the molecular basis of these phenomena and the interactions among their complex underlying pathways. Furthermore, the genomic-level information here provided on *P. chlamydosporia* should be helpful to enhance the capabilities of this fungus as a biocontrol agent of plant-parasitic nematodes and a growth-promoting agent of crop plants (Lahrmann and Zuccaro 2012).

**Acknowledgements**

This research was funded by the Spanish Ministry of Science and Innovation Grant AGL2011-29297. We acknowledge the assistance of Dax Torti, Lawrence Heisler and Tanja Durbic of the Donnelly Sequencing Center at the University of Toronto. We thank Natalia Navarro Román, (Department of Ecology, University of Alicante) as well as Rafael López and Berenice Güerri (Glen Biotech) for their help with the artwork.

**References**

Abad, P., Gouzy, J., Aury, J.M., Castagnone-Sereno P., Danchin E.G., et al., 2008. Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. Nat. Biotechnol. 26, 909-915. doi: 10.1038/nbt.1482.

Biely, P., 2012. Microbial carbohydrate esterases deacetylating plant polysaccharides. Biotechnol. Adv. 30, 1575-1588. doi: 10.1016/j.biotechadv.2012.04.010.

Bird, A.F., McClure, M.A., 1976. The tylenchid (Nematoda) egg shell: structure, composition and permeability. Parasitology. 72, 19–28. doi:10.1017/S0031182000043158.

Bordallo, J.J., Lopez-Llorca, L.V., Jansson, H-B., Salinas, J., Persmark, L., Asensio, L., 2002. Colonization of plant roots by egg-parasitic and nematode-trapping fungi. New Phytol. 154, 491–499. doi:10.1046/j.1469-8137.2002.00399.x.

Barluenga, S., Fontaine, J.G., Wang, C., Aouadi, K., Chen, R., Beebe, K., Neckers, L., Winssinger, N., 2009. Inhibition of HSP90 with pochoximes: SAR and structure-based insights. Chem. Biochem. 10, 2753-2759. doi: 10.1002/cbic.200900494.

Casas-Flores, S., Herrera-Estrella, A., 2007. Antagonism of plant parasitic nematodes by fungi. In: Kubicek, C.P., Druzhinina, I.S. (Eds.), The Mycota. Springer, Berlin, pp. 147–157. doi:10.1007/978-3-540-71840-6_9.

Cox, M.P., Peterson, D.A., Biggs, P.J., 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics 11, 485. doi: 10.1186/1471-2105-11-485.

DeZwaan, T.M., Carroll, A.M., Valent, B., Sweigard, J.A., 1999. *Magnaporthe grisea* pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive substrate cues. Plant Cell 11, 2013-2030. doi:10.1105/tpc.11.10.2013.

Dias-Arieira, C.R., Santana, S.M., de Freitas, L.G., da Cunha, T.P.L., Biela, F., Puerari, H.H., Chiamolera, F.M., 2011. Efficiency of *Pochonia chlamydosporia* in *Meloidogyne incognita* control in lettuce crop (*Lactuca sativa* L.). J. Food Agric. Environ. 9, 561-563.

Domsch, K.H., Gams, W., Anderson, T-H., 1993. Compendium of Soil Fungi. Vol I, pp. 834-836. (Eds.), IHW-Verlag, Regensburg. Germany.

Ebadi, M., Fatemy, S., Riahi, H., 2009. Evaluation of *Pochonia chlamydosporia* var. chlamydosporia as a control agent of *Meloidogyne javanica* on pistachio. Biocontrol Sci. Technol. 19, 689-700. doi:10.1080/09583150903051844.

Escudero, N., Lopez-Llorca, L.V., 2012. Effects on plant growth and root-knot nematode infection of an endophytic GFP transformant of the nematophagous fungus *Pochonia chlamydosporia*. Symbiosis 57, 33-42. doi: 10.1007/s13199-012-0173-3.

Fang W, Pava-Ripoll, M., Wang, S., St Leger, R., 2009. Protein kinase A regulates production of virulence determinants by the entomopathogenic fungus, *Metarhizium robertsii*. Fungal Genet. Biol. 46, 277-285. doi: 10.1016/j.fgb.2008.

Fang, W., St Leger, R.J., 2010. Mrt, a gene unique to fungi, encodes an oligosaccharide transporter and facilitates rhizosphere competency in *Metarhizium robertsii*. Plant Physiol. 154, 1549-1557. doi: 10.1104/pp.110.163014.

Gao, Q., Jin, K., Ying, S.H., Zhang, Y., Xiao, G., Shang, Y., Duan, Z., Hu, X., Xie, X.Q., Zhou, G., et al., 2011. Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. PLoS Genet. 7(1), e1001264.

Gibson, L.J., 2012. The hierarchical structure and mechanics of plant materials. J. R. Soc. Interface. 9, 2749-2766. doi:10.1098/rsif.2012.0341.

Gortari, M.C., Hours, R.A., 2008. Fungal chitinases and their biological role in the antagonism onto nematode eggs. A review. Mycol. Prog. 7, 221-238. doi: 10.1007/s11557-008-0571-3.

Hellwig, V., Mayer-Bartschmid, A., Müller, H., Greif, G., Kleymann, G., Zitzmann, W., Tichy, H-V., Stadler, M., 2003. Pochonins A-F, new antiviral and antiparasitic resorcylic acid lactones from *Pochonia chlamydosporia* var. catenulata. J. Nat. Prod. 66, 829-837. doi: 10.1021/np020556v.

Holt, C., Yandell, M., 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 22, 12:491.

Huang, X., Zhao, N., Zhang, K., 2004. Extracellular enzymes serving as virulence factors in nematophagous fungi involved in infection of the host. Res. Microbiol. 115, 811-816. doi: http://dx.doi.org/10.1016/j.resmic.2004.07.003.

Hung, C.Y., Seshan, K.R., Yu, J.J., Schaller, R., Xue, J., Basrur, V., Gardner, M.J., Cole, G.T. 2005. A metalloproteinase of *Coccidioides posadasii* contributes to evasion of host detection. Infect. Immun. 73, 6689-703. doi:10.1128/IAI.73.10.6689-6703.2005.

Islam, M.S., Haque, M.S., Islam, M.M., Emdad, E.M., Halim, A., Hossen, Q.M., Hossain, M.Z., Ahmed, B., et al., 2012. Tools to kill: Genome of one of the most destructive plant pathogenic fungi *Macrophomina phaseolina*. BMC Genomics 13, e493. doi: 10.1186/1471-2164-13-493.

Jackman, S.D., Birol, I., 2010. Assembling genomes using short-read sequencing technology. Genome Biol. 11, e202. doi: 10.1186/gb-2010-11-1-202.

Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P., Valent, B., 2000. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. EMBO J. 19, 4004-4014. doi:10.1093/emboj/19.15.4004.

Johnson, R., Voisey, C., Johnson, L., Pratt, J., Fleetwood, D., Khan, A., Bryan, G., 2007. Distribution of NRPS gene families within the *Neotyphodium/Epichloë* complex. Fung. Genet. Biol. 44, 1180-1190. doi:10.1016/j.fgb.2007.04.009.

Kerry, B.R., 2000. Rhizosphere interactions and the exploitation of microbial agents for the biological control of plant-parasitic nematodes. Annu. Rev. Phytopathol. 38, 423–441. doi:10.1146/annurev.phyto.38.1.423.

Kück, P., Meusemann, K., 2010. FASconCAT: Convenient handling of data matrices. Mol. Phylogenet. Evol. 56,1115-1118. doi: 10.1016/j.ympev.2010.04.024.

Lahrmann, U., Zuccaro, A., 2012. Opprimo ergo sum--evasion and suppression in the root endophytic fungus *Piriformospora indica*. Mol. Plant. Microbe Interact. 25, 727-737. doi:10.1094/MPMI-11-11-0291.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25. doi: 10.1186/gb-2009-10-3-r25.

Larriba, E., Martín-Nieto, J., Lopez-Llorca, L.V., 2012. Gene cloning, molecular modeling, and phylogenetics of serine protease P32 and serine carboxypeptidase SCP1 from nematophagous fungi *Pochonia rubescens* and *Pochonia chlamydosporia*. Can. J. Microbiol. 58, 815-827. doi: 10.1139/W2012-054.

Lechner, M., Findeiss, S., Steiner, L., Marz, M., Stadler, P.F., Prohaska, S.J., 2011. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 12, 124. doi: 10.1186/1471-2105-12-124.

Liu, K., Tian, J., Xiang, M., Liu, X., 2012. How carnivorous fungi use three-celled constricting rings to trap nematodes. Prot. Cell. 3, 325-328. doi:10.1007/s13238-012-2031-8.

Liu, S., Dean, R.A., 1997. G protein subunit genes control growth, development, and pathogenicity of *Magnaporthe grisea*. Mol. Plant. Microbe Interact. 10, 1075-1086.

Lopez-Llorca, L.V., Robertson, W.M., 1992. Immunocytochemical localization of a 32-kDa protease from the nematophagous fungus *Verticillium suchlasporium* in infected nematode eggs. Exp. Mycol. 16, 261–267. doi:10.1016/0147-5975(92)90002-9.

Lopez-Llorca, L.V., Bordallo, J.J., Salinas, J., Monfort, E., López-Serna, M.L., 2002a. Use of light and scanning electron microscopy to examine colonisation of barley rhizosphere by the nematophagous fungus *Verticillium chlamydosporium.* Micron 33, 261–267. doi:10.1016/S0968-4328(00)00070-6.

Lopez-Llorca, L.V., Olivares-Bernabeu, C., Salinas, J., Jansson, H-B., Kolattukudy, P.E., 2002b. Pre-penetration events in fungal parasitism of nematode eggs. Mycol. Res. 106, 499-506. doi:10.1017/S0953756202005798.

Lopez-Llorca, L.V., Gómez-Vidal, S., Monfort, E., Larriba, E., Casado-Vela, J., Elortza, F., Jansson, H-B., Salinas, J., Martín-Nieto. J., 2010. Expression of serine proteases in egg-parasitic nematophagous fungi during barley root colonization. Fungal Genet. Biol. 47, 342-351. doi:10.1016/j.fgb.2010.01.004.

Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl. Acids Res. 25, 955–964. doi: 10.1093/nar/25.5.955.

Macia-Vicente, J.G., Jansson, H-B., Talbot, N.J., Lopez-Llorca, L.V., 2009a. Real-time PCR quantification and live-cell imaging of endophytic colonization of barley (*Hordeum vulgare*) roots by *Fusarium equiseti* and *Pochonia chlamydosporia*. New Phytol. 182, 213-228. doi: 10.1111/j.1469-8137.2008.02743.x.

Macia-Vicente, J.G., Rosso, L.C., Ciancio, A., Jansson, H-B., Lopez-Llorca, L.V., 2009b. Colonisation of barley roots by endophytic *Fusarium equiseti* and *Pochonia chlamydosporia*: effects on plant growth and disease. Ann. Appl. Biol. 155, 391–401. doi:10.1111/j.1744-7348.2009.00352.x.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. Available at: http://journal.embnet.org/index.php/embnetjournal/article/view/200.

Mi, Q., Yang, J., Ye, F., Gan, Z., Wu, C., Niu, X., Zou, C., Zhang, K-Q., 2010. Cloning and overexpression of *Pochonia chlamydosporia* chitinase gene *pcchi44*, a potential virulence factor in infection against nematodes. Process Biochem. 45, 810–814. doi:10.1016/j.procbio.2010.01.022.

Monfort, E., Lopez-Llorca, L.V., Jansson, H-B., Salinas, J., Park, J.O., Sivasithamparam, K., 2005. Colonisation of seminal roots of wheat and barley by egg-parasitic nematophagous fungi and their effects on *Gaeumannomyces graminis* var. tritici and development of root-rot. Soil Biol. Biochem. 37, 1229–1235. doi: 10.1016/j.soilbio.2004.11.019.

Monod, M., 2008. Secreted proteases from dermatophytes. Mycopathologia 166, 285-294. doi: 10.1007/s11046-008-9105-4.

Morrissey, J.P., Osbourn, A.E., 1999. Fungal resistance to plant antibiotics as a mechanism of pathogenesis. Microbiol. Mol. Biol. Rev. 63, 708-724.

Morton, C.O., Hirsch, P.R., Peberdy, J.F., Kerry, B.R., 2003. Cloning of and genetic variation in protease VCP1 from the nematophagous fungus *Pochonia chlamydosporia*. Mycol. Res. 107, 38-46. doi:10.1017/S0953756202007050.

Newport, G., Kuo, A., Flattery, A., Gill, C., Blake, J.J., Kurtz, M.B., Abruzzo, G.K., Agabian, N., 2003. Inactivation of Kex2 diminishes the virulence of *Candida albicans*. J. Biol. Chem. 278, 1713-1720. doi:10.1074/jbc.M209713200.

Niu, X.M., Wang, Y.L., Chu, Y.S., Xue, H.X., Li, N., Wei, L.X., Mo, M.H., Zhang, K.Q., 2010. Nematodetoxic aurovertin-type metabolites from a root-knot nematode parasitic fungus *Pochonia chlamydosporia.* J. Agric. Food. Chem. 58, 828-834. doi: 10.1021/jf903259n.

Notredame, C., Higgins, D.G., Heringa, J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205-217. doi:10.1006/jmbi.2000.4042.

Palma-Guerrero, J., Jansson, H-B., Salinas, J., Lopez-Llorca, L.V., 2008. Effect of chitosan on hyphal growth and spore germination of plant pathogenic and biocontrol fungi. J. Appl. Microbiol. 104, 541–553. doi: 10.1111/j.1365-2672.2007.03567.x.

Rawlings, N.D., Morton, F.R., 2008. The MEROPS batch BLAST: a tool to detect peptidases their and non-peptidase homologues in a genome. Biochimie. 90**,** 243-259. doi: 10.1016/j.biochi.2007.09.014.

Reeves, C.D., Hu, Z., Reid, R., Kealey, J.T., 2008. Genes for the biosynthesis of the fungal polyketides hypothemycin from *Hypomyces subiculosus* and radicicol from *Pochonia chlamydosporia*. Appl. Environ. Microbiol. 74, 5121-5129. doi: 10.1128/AEM.00478-08.

Rosso, L.C., Finetti-Sialer, M.M., Hirsch, P.R., Ciancio, A., Kerry, B.R., Clark, I.M., 2011. Transcriptome analysis shows differential gene expression in the saprotrophic to parasitic transition of *Pochonia chlamydosporia.* Appl. Microbiol. Biotechnol. 90, 1981-1994. doi: 10.1007/s00253-011-3282-7.

Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18, 502–504.

Shinonaga, H., Kawamura, Y., Ikeda, A., Aoki, M., Sakai, N., Fujimoto, N., Kawashima, A., 2009. The search for a hair-growth stimulant: new radicicol analogues as WNT-5A expression

inhibitors from *Pochonia chlamydosporia* var. chlamydosporia. Tetrahedron Lett. 50, 108-110. doi: 10.1016/j.tetlet.2008.10.099.

Solomon, P.S., Waters, O.D.C., Oliver, R.P., 2007. Decoding the mannitol enigma in filamentous fungi. Trends Microbiol. 15, 257-262. doi:10.1016/j.tim.2007.04.002.

Soloviev, D.A., Jawhara, S., Fonzi, W.A., Regulation of innate immune response to *Candida albicans* infections by αMβ2-Pra1p interaction. Infect. Immun. 79, 1546–1558. doi:10.1128/IAI.00650-10.

St. Leger, R., 1995. The role of cuticle-degrading proteases in fungal pathogenesis of insects. Can. J. Bot. 73, 1119-1125. doi: 10.1139/b95-367.

St. Leger, R.J., Goettel, M., Roberts, D.W., Staples, R.C., 1991. Prepenetration events during infection of host cuticle by *Metarhizium anisopliae*. J. Invertebr. Pathol. 58, 168-179. doi: 10.1016/0022-2011(91)90061-T.

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., Morgenstern, B., 2006. AUGUSTUS: *ab initio* prediction of alternative transcripts. Nucl. Acids. Res. 34, W435-W439.

Stöver, B.C., Müller, K.F., 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinformatics 11, 7. doi: 10.1186/1471-2105-11-7.

Sung, G-H., Hywel-Jones, N.L., Sung, J-M., Luangsa-ard, J.J., Shrestha, B., Spatafora, J.W., 2007. Phylogenetic classification of *Cordyceps* and the clavicipitaceous fungi. Stud. Mycol. 57, 5-59. doi: 10.3114/sim.2007.57.01.

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y.O., Borodovsky, M., 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome. Res. 18, 1979-1990. doi: 10.1101/gr.081612.108.

Tikhonov, V.E., Lopez-Llorca, L.V., Salinas, J., Jansson, H-B., 2002. Purification and characterization of chitinases from the nematophagous fungi *Verticillium chlamydosporium* and *V. suchlasporium*. Fungal. Genet. Biol. 35, 67-78. doi: 10.1006/fgbi.2001.1312.

Trapnell, C., Pachter, L., Salzberg, S.L., 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 25, 1105-1111. doi:10.1093/bioinformatics/btp120.

Tucker, S.L., Talbot, N.J., 2001. Surface attachment and pre-penetration stage development by plant pathogenic fungi. Annu. Rev. Phytopathol. 39, 385-441. doi:10.1146/annurev.phyto.39.1.385.

Van Den Brink, J., De Vries, R.P., 2011. Fungal enzyme sets for plant polysaccharide degradation. Appl. Microbiol. Biotechnol. 91, 1477-1492. doi: 10.1007/s00253-011-3473-2 .

Wang, C., Hu, G., St. Leger, R.J., 2005. Differential gene expression by Metarhizium anisopliae growing in root exudate and host (*Manduca sexta*) cuticle or hemolymph reveals mechanisms of physiological adaptation. Fungal Genet. Biol. 42, 704-718.

Ward, E., Kerry, B.R., Manzanilla-López, R.H., Mutua, G., Devonshire, J., Kimenju, J., Hirsch, P.R., 2012. The *Pochonia chlamydosporia* serine protease gene vcp1 is subject to regulation by carbon, nitrogen and pH: Implications for Nematode Biocontrol. PLoS One. 7, e35657. doi:10.1371/journal.pone.0035657.

Wharton, D., 1980. Nematode egg-shells. Parasitology. 81, 447-463.

Winnenburg, R., Urban, M., Beacham, A., Baldwin, T.K., Holland, S., Lindeberg, M., Hansen, H., Rawlings, C., Hammond-Kosack, K.E., Kohler, J., 2008. PHI-base update: additions to the pathogen host interaction database. Nucleic Acids. Res. 36, D572-D576. doi:10.1093/nar/gkm858.

Xiao, G., Ying, S.H., Zheng, P., Wang, Z.L., Zhang, S., et al., 2012. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. Sci. Rep. 2, 483. doi:10.1038/srep00483.

Yandell, M., Ence, D., 2012. A beginner's guide to eukaryotic genome annotation. Nat. Rev. Genet. 13, 329-342. doi: 10.1038/nrg3174.

Yang, J., Wang, L., Ji, X., Feng, Y., Li, X., et al., 2011. Genomic and proteomic analyses of the fungus *Arthrobotrys oligospora* provide insights into nematode-trap formation. PLoS Pathog. 7, e1002179. doi:10.1371/journal.ppat.1002179.

Yang, J.K., Tian, B.Y., Liang, L.M., Zhang, K.Q., 2007. Extracellular enzymes and the pathogenesis of nematophagous fungi. Appl. Microbiol. Biotechnol. 75, 21-31. doi: 10.1007/s00253-007-0881-4.

Zare, R., Gams, W., Evans, H. C., 2001. A revision of Verticillium section Prostrata V. The genus Pochonia, with notes on Rotiferophthora. Nova Hedwigia 73, 51–86.

Zhang, C., Xia, Y., 2008. Identification of genes preferentially expressed during microcycle conidiation of *Metarhizium anisopliae* using suppression subtractive hybridization. FEMS Microbiol. Lett. 286, 71-7. doi: 10.1111/j.1574-6968.2008.01257.x.

Zhang, C., Xia, Y., 2009. Identification of genes differentially expressed in vivo by Metarhizium anisopliae in the hemolymph of *Locusta migratoria* using suppression-subtractive hybridization. Curr. Genet. 55, 399-407. doi: 10.1007/s00294-009-0254-x.

Zhao, Y., Park, R-D., Muzzarelli, R.A.A., 2010. Chitin deacetylases: Properties and applications. Mar. Drugs. 8, 24-46. doi: 10.3390/md8010024.

Zheng, P., Xia, Y., Xiao, G., Xiong, C., Hu, X., et al., 2011. Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued traditional Chinese medicine. Genome. Biol. 12, R116. doi: 10.1186/gb-2011-12-11-r116.

Zhou, H., Qiao, K., Gao, Z., Vederas. J.C., Tang, Y., 2010. Insights into radicicol biosynthesis via heterologous synthesis of intermediates and analogs. J. Biol. Chem. 285, 41412-41421. doi: 10.1074/jbc.M110.183574.

Zuccaro, A., Lahrmann, U., Güldener, U., Langen, G., Pfiffi, S., et al., 2011. Endophytic life strategies decoded by genome and transcriptome analyses of the mutualistic root symbiont *Piriformospora indica*. PLoS Pathog. 7, e1002290. doi 10.1371/journal.ppat.1002290.

**Figures**.

Fig. 1. **Phylogenomic tree of *P. chlamydosporia*.** A maximum likelihood phylogenomic tree was constructed on the basis of 382 concatenated orthologous protein-coding genes using the Dayhoff amino acid substitution model showing the evolutionary relationships between the indicated 29 fungal species. Bootstrap values are shown on relevant tree branches.

Fig. 2. **Mobile elements predicted in the genomes of *P. chlamydosporia* and other fungal pathogens of invertebrates.** A) Mobile elements classified into subclasses predicted in the genome. B) Families of mobile elements predicted in the sequenced genomes of fungal pathogens of invertebrates. Class II transposons are in yellow, class I retrotransposons in blue and unknown elements in green. The families of mobile elements in the *P. chlamydosporia* genome are indicated by brackets. Subclasses and families of mobile elements are indicated below each bar.

Fig. 3. **Gene Ontology (GO) functional annotation of proteins encoded by genes in the *P. chlamydosporia* genome.** GO charts were generated using generic Slim terms at level 3 within the GO Biological Process (left) and Molecular Function (right) domains. A) Protein-coding genes predicted in the *P. chlamydosporia* genome. B) *P. chlamydosporia* genes with homologous counterparts in the PHI database. C) Genes expressed during endophytic colonization of barley roots by *P. chlamydosporia* detected by RNA-Seq.

Fig. 4. **Families of proteins encoded in the *P. chlamydosporia* genome.** Protein families with an involvement in the different P. chlamydosporia lifestyles are shown grouped by different colors according to their biological roles. Each bar represents the number of members of each protein family predicted in the genome, and the stripped area indicates the number of them predicted to be extracellular. The black line indicates the number of genes within each family that are expressed during endophytic root colonization.

Fig. 5. **Hydrolytic enzymes encoded by the *P. chlamydosporia* genome and expressed during endophytic behavior.** *A) Sector graph showing the percentage of hydrolytic enzyme-*

*coding genes identified within each catalytic type that are expressed during root endophytic colonization. B) Selected hydrolytic enzymes are grouped by different colors according to their substrate types. Each bar represents the number of hydrolytic enzymes belonging to each family represented in the genome, and the stripped area the number of them predicted to be secreted. The red line indicates the number of genes within each hydrolase family expressed during endophytic colonization.*

Fig. 6. **Tritrophic lifestyles of the nematophagous fungus *P. chlamydosporia*.** The three trophic modes of *P. chlamydosporia* as nematode-egg parasite (P), soil saprophyte (S) and root endophyte (E), are shown in the figure. The parasitic life-mode of *P. chlamydosporia* is illustrated (P inset) by an image of a GFP transformant strain of this fungus infecting a plant-parasitic nematode (*Meloidogyne javanica*) egg. The fungus forms an appresorium on the egg-shell from which penetration and colonization of egg contents takes place (from Escudero and Lopez-Llorca, 2012). *P. chlamydosporia* is a true endophyte, since it colonizes living barley root cells (E inset). The image portraits the fungus penetrating the cell wall of two adjacent cortex cells(arrowheads). The sample has been labeled with the membrane tracker FM4-64 to show that the fungus-colonized cells retain membrane integrity (blue staining) (from Macia-Vicente et al., 2009) Finally, the saprophytic lifestyle of *P. chlamydosporia* in the soil is illustrated (S inset) by a chlamydospore (resting stage) of the fungus (picture gift from Dr. Palma-Guerrero, U. Berkeley, USA). *P. chlamydosporia* switches lifestyles and makes a differential use of protein families encoded in its genome. The suggested importance of key protein families in the different lifestyles of this fungus is indicated by the font size of its corresponding protein family abbreviation. Hydrolytic enzymes are in brackets. Abbreviations: PROT: proteases, GH: glycoside hydrolases, CE: carbohydrate esterases, DEX: detoxification, SM: secondary metabolism, TRS, transporters, TF: transcription factors, ST: signal transducers.

**Table 1. Main features of the *Pochonia chlamydosporia* genome**

| | |
|---|---|
| Genome size (Mb) | 41.2 |
| Assembly N50 (kb) | 225 |
| Coverage (fold) | 136 |
| G+C content (%) | 49.9 |
| Repeat rate (%) | 0.46 |
| Protein-coding genes | 12,122 |
| Unique proteins | 672 |
| Secreted proteins | 2,485 |
| Transmembrane proteins | 2,707 |
| Genome coding (%) | 42.1 |
| Gene density (genes/Mb) | 294 |

| | |
|---|---|
| Average ORF length (kb) | 1.4 |
| tRNA genes | 45 |
| Mobile elements | 349 |

***Table 2. Genome size, number of predicted protein-coding genes, G+C content and best Blast top hit homology between the genomes of P. chlamydosporia and those of 28 other fungi.***

| Species | Genome size (Mb)[a] | G+C (%)[a] | Number of proteins[a] | BBTH[b] | Lifestyle | NCBI Bioproject |
|---|---|---|---|---|---|---|

| | | | | | | |
|---|---|---|---|---|---|---|
| *Pochonia chlamydosporia* | 41.2 | 49.9 | 12,122 | – | Nematophagous | PRJNA68669 |
| *Metarhizium anisopliae* | 39.0 | 51.5 | 10,582 | 4,169 | Entomopathogenic | PRJNA38717 |
| *Metarhizium acridum* | 38.0 | 49.9 | 9,849 | 3,608 | Entomopathogenic | PRJNA38715 |
| *Epichloë festucae* | 34.7 | 44.3 | 9,273 | 871 | Endophytic | PRJNA42133 |
| *Fusarium solani* | 51.2 | 50.8 | 15,708 | 411 | Phytopathogenic | PRJNA51499 |
| *Trichoderma virens* | 39.0 | 49.2 | 12,406 | 386 | Mycopathogenic | PRJNA19983 |
| *Fusarium oxysporum* | 61.4 | 48.2 | 15,438 | 271 | Phytopathogenic | PRJNA174274 |
| *Trichoderma atroviride* | 36.1 | 49.7 | 11,816 | 227 | Mycopathogenic | PRJNA19867 |
| *Beauveria bassiana* | 33.6 | 51.5 | 10,364 | 187 | Entomopathogenic | PRJNA38719 |
| *Cordyceps militaris* | 32.2 | 51.4 | 9,651 | 162 | Entomopathogenic | PRJNA41129 |
| *Trichoderma reesei* | 33.3 | 52.8 | 9,115 | 140 | Mycopathogenic | PRJNA15571 |
| *Aspergillus oryzae* | 36.5 | 48.3 | 11,397 | 124 | Saprophytic | PRJNA88495 |
| *Fusarium graminearum* | 72.7 | 48.3 | 11,628 | 111 | Phytopathogenic | PRJNA243 |
| *Colletotrichum graminicola* | 50.9 | 49.1 | 12,02 | 99 | Phytopathogenic | PRJNA37879 |
| *Aspergillus niger* | 35.7 | 50.4 | 11,182 | 92 | Saprophytic | PRJNA19263 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Verticillium dahliae* | 32.9 | 55.8 | 10,535 | 86 | Phytopathogenic | PRJNA28529 |
| *Colletotrichum higginsianum* | 44.1 | 55.1 | 16,141 | 83 | Phytopathogenic | PRJNA47061 |
| *Botryotinia fuckeliana* | 39.5 | 43.5 | 16,389 | 72 | Phytopathogenic | PRJNA20061 |
| *Aspergillus clavatus* | 27.8 | 49.2 | 9,121 | 62 | Saprophytic | PRJNA18467 |
| *Arthrobotrys oligospora* | 39.9 | 44.5 | 11,479 | 53 | Nematophagous | PRJNA41495 |
| *Magnaporthe oryzae* | 40.9 | 51.6 | 12,836 | 47 | Phytopathogenic | PRJNA13840 |
| *Verticillium albo-atrum* | 30.3 | 56 | 10,237 | 45 | Phytopathogenic | PRJNA51263 |
| *Neurospora crassa* | 39.2 | 49.3 | 9,841 | 31 | Saprophytic | PRJNA132 |
| *Laccaria bicolor* | 58.7 | 47 | 18,215 | 10 | Mycorrhizal | PRJNA29019 |
| *Piriformospora indica* | 25.0 | 50.6 | 11,791 | 6 | Endophytic | PRJEA76339 |
| *Saccharomyces cerevisiae* | 12.1 | 38.2 | 5,905 | 6 | Saprophytic | PRJNA128 |
| *Ustilago maydis* | 19.9 | 53.7 | 6,548 | 5 | Phytopathogenic | PRJNA14007 |
| *Schizosaccharomyces pombe* | 12.5 | 36 | 5,133 | 2 | Saprophytic | PRJNA127 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Kluyveromyces lactis* | 10.7 | 38.7 | 5,076 | 1 | Saprophytic | PRJNA13835 |

[a]Genome size, G+C content and number of protein-coding genes for each fungus were obtained from the Genome database at NCBI.

[b]BBTH (Best Blast Top Hit) refers to the number of sequences returning the top hit after a BlastP search of 28 fungal genomes.

Figure 1



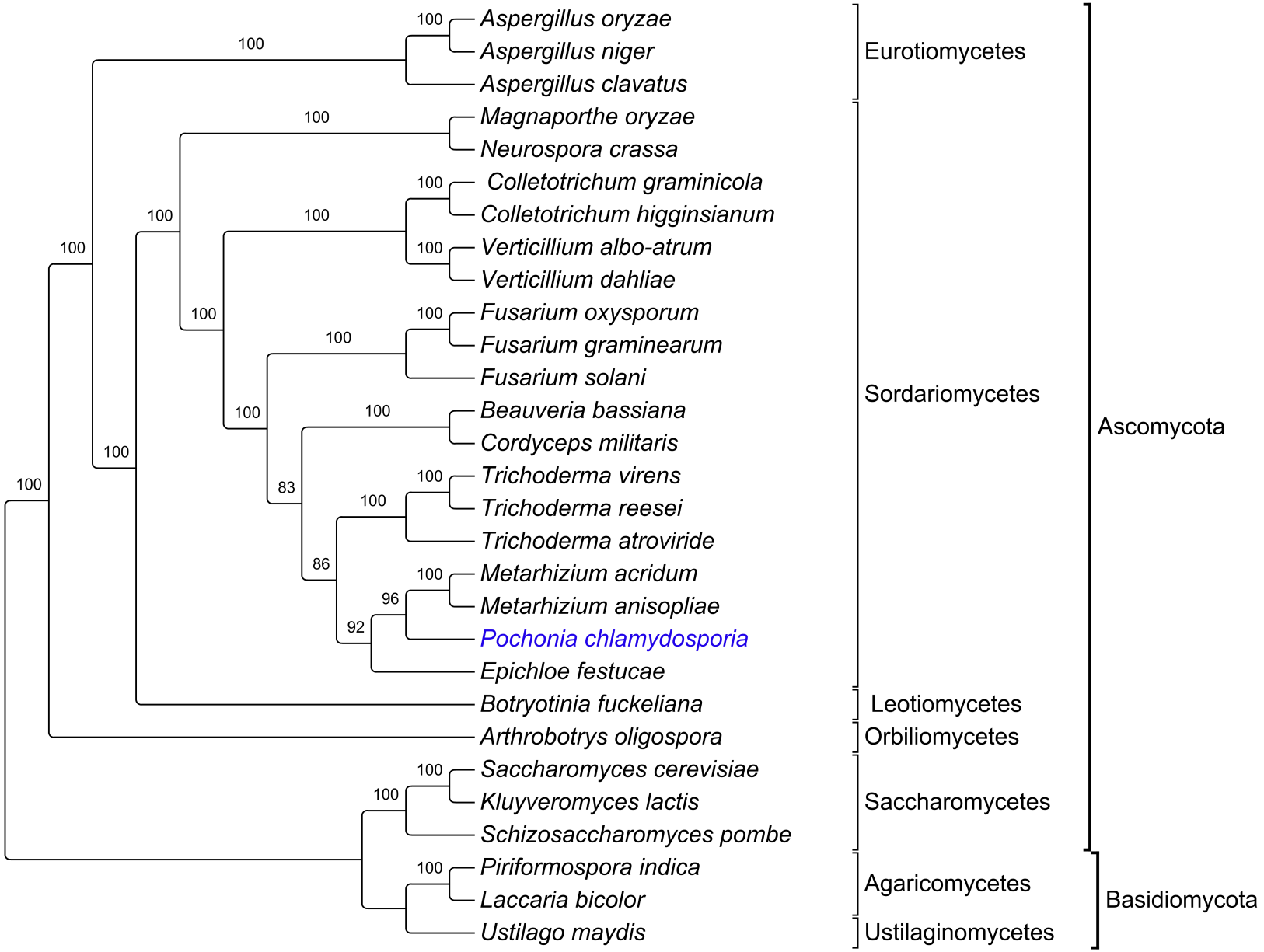

Figure 2



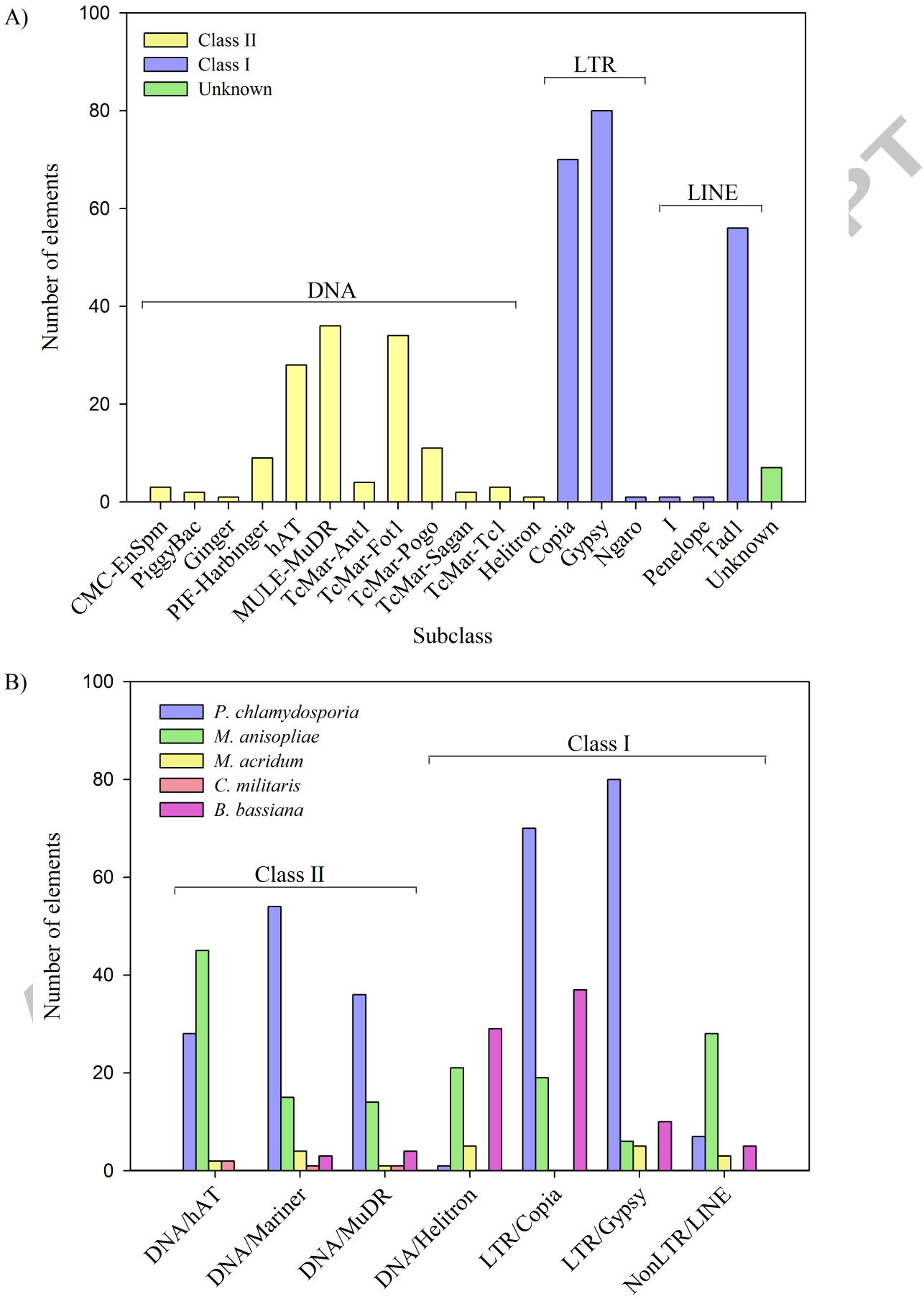

Figure 3



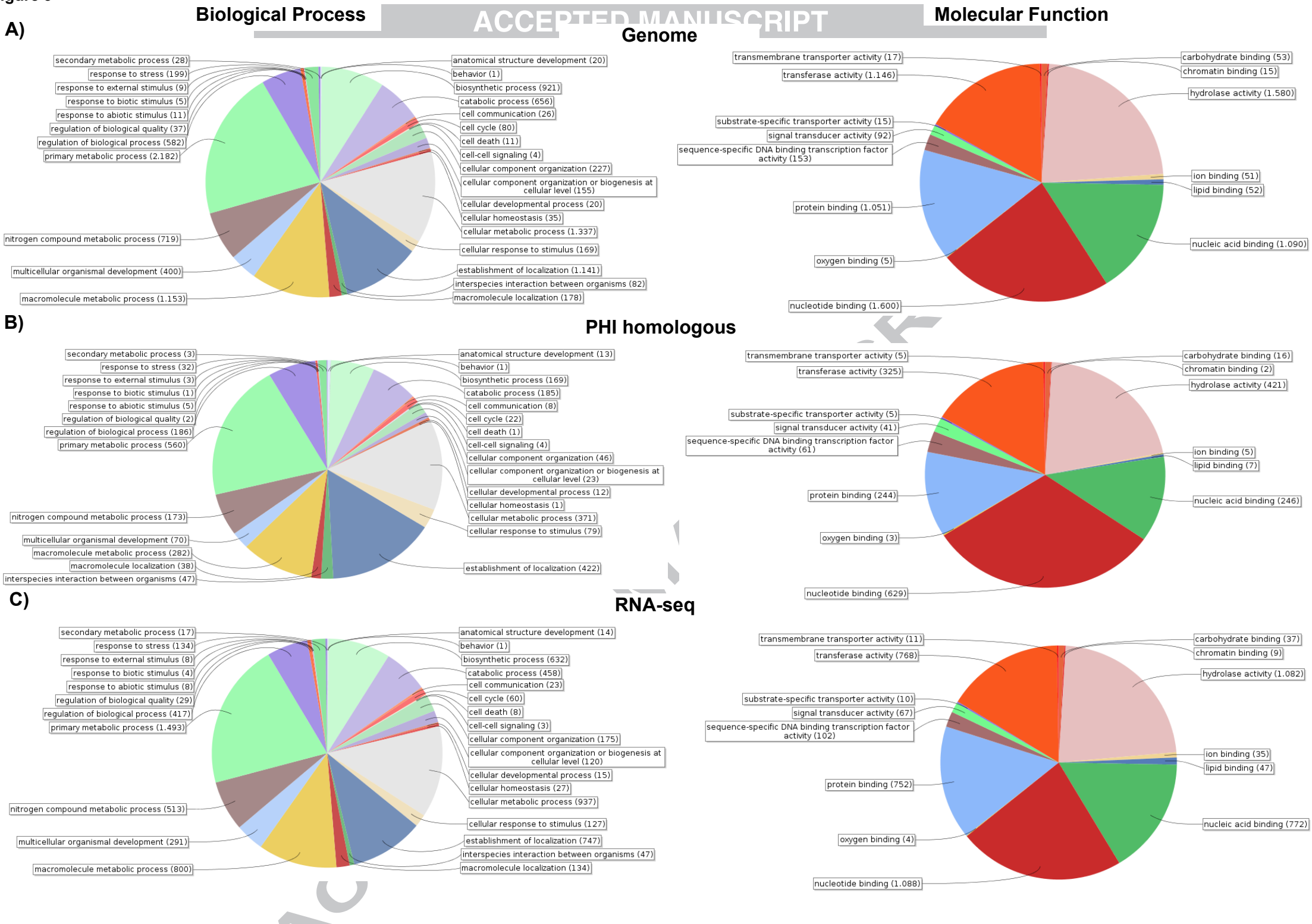

Figure 4



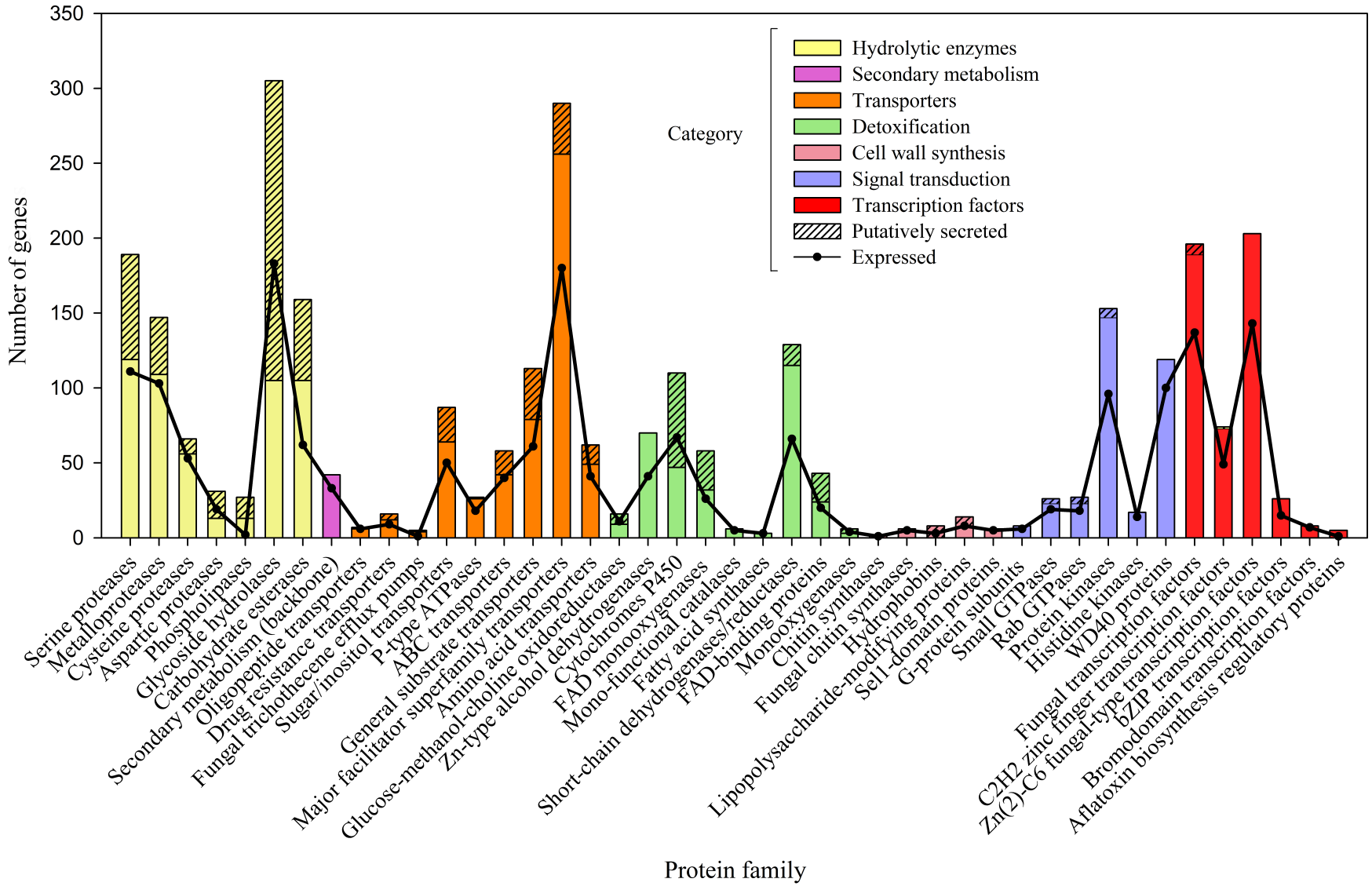

Figure 5



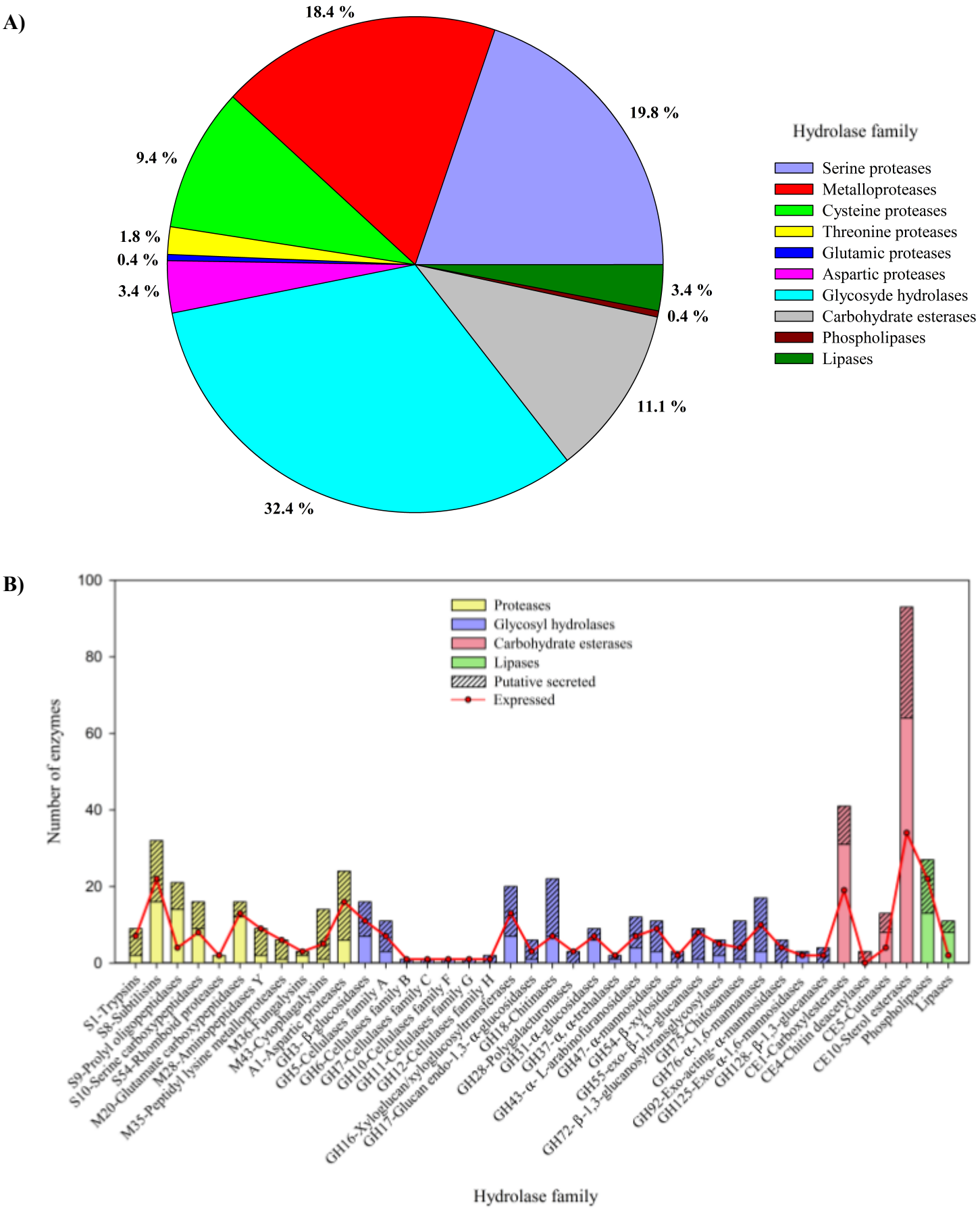

Figure 6