FUNGAL BIOLOGY REVIEWS 33 (2019) 82–97

British Mycological Society promoting fungal science

journal homepage: www.elsevier.com/locate/fbr

Review

Origin of fungal biomass degrading enzymes: Evolution, diversity and function of enzymes of early lineage fungi

Lene LANGE[a,*], *Bo PILGAARD*[b], *Florian-Alexander HERBST*[c], *Peter Kamp BUSK*[d], *Frank GLEASON*[e], *Anders Gorm PEDERSEN*[f]

[a]*Center for Bioprocess Engineering, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark*
[b]*Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark*
[c]*Department of Chemistry and Bioscience Aalborg University, Denmark*
[d]*Department of Science and Environment, Roskilde University, Denmark*
[e]*University of Sydney, Sydney, NSW, 2006, Australia*
[f]*Department of Bio and Health Informatics, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark*

ARTICLE INFO

Article history:
Received 6 June 2018
Received in revised form
31 August 2018
Accepted 5 September 2018

Keywords:
AA11
AA9
Early lineage fungi
Enzyme phylogeny
Enzyme secretome
Evolution
LPMO
Rhizophlyctis rosea

ABSTRACT

The aim of this study was to elucidate the evolution of enzyme secretome of early lineage fungi to contribute to resolving the basal part of Fungal Kingdom and pave the way for industrial evaluation of their unique enzymes. By combining results of advanced sequence analysis with secretome mass spectrometry and phylogenetic trees, we provide evidence for that plant cell wall degrading enzymes of higher fungi share a common ancestor with enzymes from aerobic ancient fungi. Sequence analysis (HotPep, confirmed by dbCAN-HMM models) enabled prediction of enzyme function directly from sequence. For the first time, oxidative enzymes are described here in early lineage fungi (Chytridiomycota & Cryptomycota), which supports the conceptually new understanding that fungal LPMOs were also present in the early evolution of the Fungal Kingdom. Phylogenetic analysis of fungal AA9 proteins suggests an LPMO-common-ancestor with Ascomycetes and Basidiomycetes and describes a new clade of AA9s. We identified two very strong biomass degraders, *Rhizophlyctis rosea* (soil-inhabiting) and *Neocallimastix californiae* (rumen), with a rich spectrum of cellulolytic, xylanolytic and pectinolytic enzymes, characteristically including several different enzymes with the same function. Their secretome composition suggests horizontal gene transfer was involved in transition to terrestrial and rumen habitats. Methods developed for recombinant production and protein characterization of enzymes from zoosporic fungi pave the way for biotechnological exploitation of unique enzymes from early lineage fungi with potential to contribute to improved biomass

* *Corresponding author.*
E-mail address: lene.lange2@gmail.com (L. Lange).
https://doi.org/10.1016/j.fbr.2018.09.001

conversion. The phyla of ancient fungi through evolution have developed to be very different and together they constitute a rich enzyme discovery pool.

1. Introduction

The aim of this study was to elucidate the enzyme secretome composition of zoosporic early lineage fungi, across all phyla (James *et al.*, 2006b; Kämper *et al.*, 2006; Powell and Letcher, 2014); to contribute to resolving the basal part of the fungal tree of life, and to pave the way for industrial use of their unique enzymes by developing methods for cloning, expression and protein characterization of enzymes from such fungi.

Zoosporic fungi have been described as having widespread ecological importance for global plant biomass conversion due to their role in recycling global carbon (Chang *et al.*, 2015; Gleason *et al.*, 2017; Jephcott *et al.*, 2016). Yet early lineage fungi remain underexplored, especially at the molecular level. While evolution and phylogeny of the higher Dikarya fungi have been elucidated (Ruggiero *et al.*, 2015) the basal lineages of the Fungal Tree of Life remain largely unresolved. Contributions to resolving its evolution will have to come from advanced sequence analysis and phylogeny (Chang *et al.*, 2015; Taylor and Berbee, 2006).

James *et al.*, (2006a) advanced understanding of the early evolution of fungi by the use of a six-gene phylogeny that enabled identification of the last common ancestor of crown group fungi. James *et al.*, (2006b) described Blastocladiomycota as a new phylum. Hibbett *et al.*, (2007) assembled a higher level classification of fungi, including the early lineages. They retained the phylum Chytridiomycota and proposed a new phylum, Neocallimastigomycota. Among the early lineage zoosporic fungi, *Olpidium* and *Rozella* long remained unclassified (Hibbett *et al.*, 2007). Jones *et al.*, (2011) described Cryptomycota as a new phylum, and James *et al.*, (2013) analyzed the genome of *Rozella allomycis* - the first species of this new phylum to be genome sequenced. Based on morphological phenotypic characters, zoospore ultrastructure and DNA-based molecular methods Powell and Letcher, (2012) recognized the following phyla: Chytridiomycota, Monoblepharidomycota and Neocallimastigomycota. However, Ruggiero *et al.*, (2015) grouped all zoosporic fungi into one phylum, and ranked the anaerobic rumen fungi only as a class. The paper by Chang *et al.*, (2015) marked a breakthrough with regard to including enzymes in descriptions of classification and evolution of the early lineage fungi. Using expansion of pectinase genes and the phylogeny of the organism, Chang *et al.* (l.c.) estimated the genomic time clock of the plant–fungus association. Such advanced studies were taken further by the work of the Spatafora and Berbee research groups (Berbee *et al.*, 2017; Spatafora *et al.*, 2016).

In the current study, we focus on the molecular, functional and structural diversity of enzymes from early lineage fungi as key characteristics for their biology, ecology and evolution. The composition of the fungal secretome metabolizes substrates to provide nutrients, and is therefore a determining factor for competitiveness of the fungus, thus vital for growth and reproduction. However, the findings of Krijger *et al.* (2014) support the observation that the fungal secretome is influenced more by phylogeny than lifestyle.

The fast track methodological approach, used in the current study to elucidate the CAZy enzyme profiles (Cantarel *et al.*, 2009) builds on a combined use of the non-alignment based HotPep (Homology-to-Peptides) sequence analysis program (Busk *et al.*, 2017) and a dbCAN analysis (Huang *et al.*, 2018; Yin *et al.*, 2012). The basis for such analysis is the well-curated CAZy database, which groups enzymes into enzyme families (and subfamilies). Furthermore, HotPep can capture protein features of importance for enzyme function; mass spectrometry studies identify presence of proteins in the secretome; and HotPep-mediated prediction of function (going beyond annotation to enzyme family) enables biological and evolutionary interpretations across organismal habitats and physiological and taxonomic groupings.

2. Materials and methodologies for comparative secretome studies

Rhizophlyctis rosea (de Bary and Woronin) Fischer, NRBC 105426, was isolated from soil collected in Sydney University garden using filter paper as bait (Gleason *et al.*, 2004) and deposited in the NITE Biological Resource Center (see Fig. 1, Supplementary Fig. A.1). The medium used for inducing production of plant cell wall degrading enzymes consisted of 1 % Avicel (Sigma), 0.5 % CMC (Sigma), 0.5 % wheat bran-broth (Finax, Denmark), 0,5 % yeast extract (Sigma), 0.1 % mycological peptone (Sigma) and 0.05 % glucose (Sigma). For full description of methods, see Supplementary Text A.1.

Eleven different AZCL substrates (Megazyme, Ireland) were prepared using Britten Robinson buffer. The plates were incubated at 30 °C overnight and the radius of the zones measured. Endoglucanase activity was measured as described earlier (Busk and Lange, 2013a).

DNA was extracted as described (Busk and Pagès, 2002) and sequenced twice on an Illumina Hiseq 2000 in one multiplexed lane as Paired-End libraries using Truseq chemistry. The genome coverage of the two PE-libraries was estimated to 140x and 279x. Before assembly, the raw reads were assessed with FastQC v0.10.1 (Andrews, 2010) and processed according to (Lindgreen, 2012), (Martin, 2011) and (Kelley *et al.*, 2010) and assembled with Velvet 1.2.10 (Zerbino and Birney, 2008) The genome was submitted to Genbank (https://www.ncbi.nlm.nih.gov/genbank/) and given the accession number GCA002214945.1. For GenBank submissions, see Supplementary Tab. A.1.

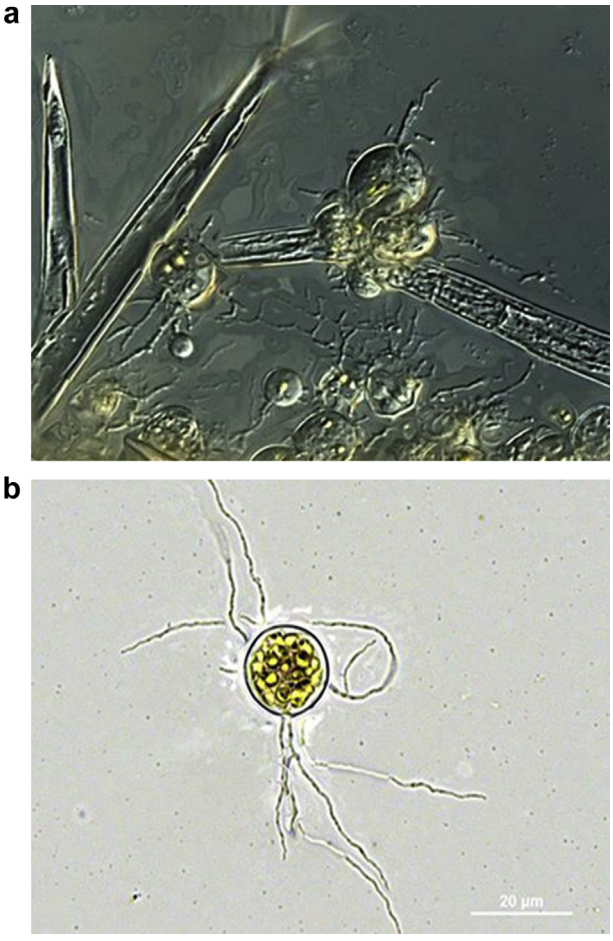

Fig. 1 – a,b Light micrographs of *Rhizophlyctis rosea*, NRBC 105426. In laboratory cultures, the rhizoids are thinner and the color fainter than when observed directly from new samplings from agricultural fields. In (a), the *R. rosea* rhizoids can be seen in intimate contact with the cellulose fibers of the lignocellulosic husk (Courtesy of Linda Henderson, The University of Sydney).

All annotations were performed on whole genome predicted protein sequences. For *R. rosea* and *H. polyrhiza* AUGUSTUS 2.7 (Stanke *et al.*, 2008) was utilized with *R. oryzae* and *B. dendrobatitis* as model organisms for gene predictions from the whole genome. PPR (Busk and Lange, 2013b) based HotPep (Busk *et al.*, 2017) and dbCAN (Huang *et al.*, 2018) based HMMer3 (Eddy, 2011) were used for annotation and functional prediction of carbohydrate active enzymes. Only hits confirmed by both approaches were used for genome wide comparisons. All protein sequences submitted to Genbank from *R. rosea* (Supplementary Tab. A.1) were manually inspected. Signal peptides were predicted using the SignalP 4.1 (Petersen *et al.*, 2011) and Phobius (Käll *et al.*, 2007) servers.

Total RNA was extracted and purified using the RNeasy Plant Mini Kit (Qiagen, Germany). cDNA was synthesized with oligodT random primer and ImProm II reverse transcriptase (Promega, USA) as described (Madsen *et al.*, 2009). The synthesized cDNA was used as template for the subsequent PCRs (Tab. A2).

Protein expression was performed using the *PichiaPink*™ Expression System protocol (Invitrogen, USA) (see Supplementary Fig. A.2).

Preparing for LC-MS/MS: The total protein content from the triplicate supernatants were precipitated by trichloroacetic acid (Sigma, Germany) overnight at −20 °C. The protein pellet was solubilized in digestion buffer (1 % sodium deoxycholate, 50 mM triethylammonium bicarbonate, pH 8.0) and subsequently reduced, alkylated, in-solution digested, desalted and measured by LC-MS/MS as previously described (Huang *et al.*, 2015). Protein identification was done using the open-source software MaxQuant v1.5.2.8 (Tyanova *et al.*, 2016).

PCA analysis was performed and visualized in Rstudio 1.1.383 (RStudio Team, 2016) using R version 3.4.3 (R core team, 2017) and the factoextra package (Kassambara and Mundt, 2017) with the scaled family annotations for all the genomes as dataset.

For construction of phylogenetic trees, publically available protein sequences were obtained from Genbank. The catalytic domains of these sequences used for phylogeny were defined using the NCBI CDD batch search (Marchler-Bauer *et al.*, 2015). For full list of accession numbers, see Supplementary Tab. A.3. Multiple alignments were performed with MAFFT version 7 (Katoh and Toh, 2008) using the E-INS-i method. Phylogenies were reconstructed with Bayesian inference using MrBayes v3.2.4 that averaged over 10 aa substitution models and assumed gamma-distributed rates over sites. The runs continued until convergence was reached across 2 sets of 3 chains: "R-hat" values for parameters near 1 and standard deviation of split frequencies below 0.01. The trees were visualized with Fig-Tree v1.4.2 (http://tree.bio.ed.ac.uk/software/Figtree/).

3. Results

The results from re-analyzing published genome data, put in perspective by new information, are grouped into four parts, 3.1–3.4.

3.1. R. rosea *as a model species for studying enzyme secretome of early lineage fungi*

The soil inhabiting chytrid, *R. rosea* (Chytridiomycota), was chosen for studying plant biomass degrading enzymes among the early lineage zoosporic fungi. The aim was to develop methods for cloning, expression and characterization of enzymes, to facilitate investigation of evolution of biotroph and saprotroph roles of such fungi in nature, and at the same time to pave the way for industrial use of the enzymes from this hitherto unexploited group of microorganisms.

The *R. rosea* isolate NRBC 105426 (Gleason *et al.*, 2004; Pilgaard, 2014) was selected based on its activity profile characterized by testing the culture broth when grown under inducing conditions (see Table 1). The light micrographs (Fig. 1) show the morphology of the *R. rosea* isolate studied when grown in lab-culture. Genome sequencing of *R. rosea* was performed as paired-end sequencing on an Illumina Hiseq 2000. The draft genome assembly yielded 4476 contigs with sizes ranging from 200 b to 309 kb and a median size of 12.4 kb; total length of assembly approx. 47 Mb.

The single copy gene GH45 (predicted by HotPep analysis to have endo-1,4-β-D-glucanase, EC 3.2.1.4 function) was selected for cloning. It was recombinantly produced in *Pichia pastoris*. The predicted gene RrCel45 was 104 bp and 897 bp long with and without two introns (lengths confirmed by sequencing PCR products of the cDNA and gDNA). A CDD search confirmed the GH45 domain and revealed a CBM1 in N-terminal position. RrCel45 was successfully expressed in *P. pastoris*. MALDI-TOF analysis of the SDS-band confirmed the identity of the protein; and the purity of the protein was ≥95 % (see Fig. A2). Two more enzymes from *R. rosea* (β-xylosidase, GH43, and endo-xylanase, GH11) were successfully expressed in *P. pastoris* (to be reported separately by Huang, Lange *et al.*).

The RrCel45 activity was tested on azurine crosslinked substrates. This revealed high activity towards carboxymethyl cellulose and less activity towards α -cellulose and β-glucan. No activity was found on crystalline cellulose or xylan. Optimum pH was pH5 at 35 °C; and optimum temperature was 35 °C at pH5 (Fig. 2). Thermostability analysis of RrCel45 revealed high thermostability, retaining approx. 40 % residual activity after incubation at 100 °C for 1 h. The specific activity of RrCel45 was estimated to be low (0.5 U/mg of protein).

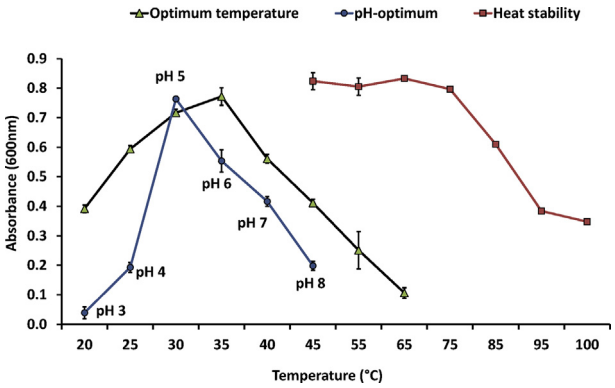

Fig. 2 – Characterization of optimum temperature and pH indicate a typical fungal enzyme profile. However, temperature stability was exceptionally high: Stable up to 70C; 40 % residual activity of intact protein, active at 100C.

3.2. *Elucidating enzyme secretome composition of the earliest branches of the fungal tree of life*

One thousand two hundred and twenty eight proteins were found by aligning MS data (from *R. rosea* culture broth) to the genome predicted proteins. Of these, 124 proteins were recognized by HotPep/dbCAN analysis to belong to four CAZy enzyme categories, Auxiliary Activity (AA), Polysaccharide Lyases (PL), Carbohydrate Esterases (CE) and Glycohydrolases (GH). Domain searches resulted in 46 additional proteins recognized as putative carbohydrate active enzymes. Supplementary Tab. A.4 gives a comparison of carbohydrate active enzymes of genome and secretome of *R. rosea*. Table 2 lists the top 20 % of the most abundant carbohydrate active enzymes in *R. rosea* secretome predicted by HotPep/dbCAN analysis. The 20 % top scoring enzymes with regard to relative secretome abundance were GH6, GH7, GH10 and GH11. Abundance was calculated as a relative value based on the total amount of predicted carbohydrate active enzymes in the MS analysis. The MS-based estimate of enzyme protein abundance, summed up for each type of function, is presented in Table 3. These data show that the three most dominating functions found in the secretome of *R. rosea*, ranked from one to three, were endo 1,4 β xylanase (EC 3.2.1.8), LPMO AA9, and endo 1,4 β glucanase (EC 3.2.1.4). The high score of AA9 is found by summing up the 20 AA9 proteins found by MS analysis of the secretome.

Available sequencing data from 11 published zoosporic early lineage species (Supplementary Tab. A.5) (Chang *et al.*, 2015; Grigoriev *et al.*, 2012; James *et al.*, 2013; Joneson *et al.*, 2011; Nordberg *et al.*, 2014; Russ *et al.*, 2016; Youssef *et al.*, 2013) and the *de novo* genome sequencing data of *R. rosea*, were analyzed for carbohydrate active enzymes by HotPep/dbCAN. Only the hits confirmed by both HotPep and dbCAN are included in this study; for statistics of annotation by HotPep as compared to dbCAN, see Supplementary Tab. A.6. The 12 genomes represent all four fungal zoosporic early lineage phyla: Chytridiomycota, Neocallimastigomycota, Blastocladiomycota and Cryptomycota; see taxonomic overview, Supplementary Tab. A.7 (for taxonomic affiliation of *R. rosea*, see (Letcher *et al.*, 2008) and for *R. globosum* see (Letcher *et al.*, 2006). A list of annotated CAZy hits, assigned to protein family based on HotPep/dbCAN is shown in Supplementary Tab. A.8.

The two strongest biomass degraders were *R. rosea* and the anaerobic cow rumen fungus, *Neocallimastix californiae*. Both HotPep and dbCAN annotated these two hotspots to have a high number of predicted CAZy genes (up to 225 and 667, respectively; Supplementary Tab. A.6; and Tables 4a,b and 5a,b; and to have several genes encoding enzymes with the same e.g. cellulolytic or xylanolytic activity (Busk *et al.*, 2014). *Gonapodaya prolifera* also had a broad spectrum of CAZy enzymes, however, this species was richest not in

Table 1 – Enzyme activity profile of *R. rosea* culture supernatant with a strong profile of plant cell wall degrading enzyme activities.

Substrate	β-glucan	HE-Cellulose	Galacto-mannan	Galactan (potato)	Xyloglucan (tamarind)	Xylan (birch)	Arabino-xylan
CMC	56	50	33	28	50	50	129
Avicel	73	50	25	13	50	50	121
Mixed-medium (MS)	154	28	79	13	79	113	113

The activity profile (scored by measuring the blue halo clearing zones) was estimated by semi-quantitative AZCL-plate assays. Units calculated in mm^2.

Table 2 – Top 20 % of the most abundant carbohydrate active enzyme proteins of the *R. rosea* secretome.

Cazy family	Predicted function by EC	Abundance %	Trans-membrane domain	Signal Peptide
GH7	3.2.1.4	9.08	0	Y
GH10	3.2.1.8	6.86	0	Y
GH6	3.2.1.91	6.65	0	Y
GH7	3.2.1.176	6.57	0	0
GH11	3.2.1.8	5.10	0	Y
CE1	3.1.1.72	3.47	0	Y
CE4	3.5.1.41	3.10	1	Y
AA9	0	3.10	0	Y
CE4	3.5.1.41	2.79	1	Y
AA9	0	2.52	0	Y
GH10	3.2.1.8	2.01	0	Y
GH26	0	1.94	1	0
GH10	3.2.1.8	1.89	0	0
AA8	1.1.99.18	1.77	0	Y
AA9	0	1.68	0	Y
CE4	3.5.1.41	1.60	0	Y
GH43	3.2.1.37	1.28	0	0
GH10	3.2.1.8	0.98	0	Y
GH3	3.2.1.21	0.92	0	0
GH26	0	0.79	0	0
GH35	3.2.1.23	0.79	1	0
AA6	0	0.74	0	0
GH5	3.2.1.4	0.69	0	Y
CE1	3.1.1.72	0.59	0	0

The abundance value is a relative value based on the total amount of predicted carbohydrate active enzyme proteins in the MS analysis. The top 20 % enzymes listed are mainly involved in cellulolytic (EC 3.2.1.4, 3.2.1.91, 3.2.1.176, 3.2.1.21, AA9, AA8) and xylanolytic (EC 3.2.1.8, 3.2.1.37, 3.1.1.72) degradation.

Table 3 – MS-based estimate of enzyme protein abundance, summed up for each type of function.

Function	endo-1,4-β-xylanase	AA9 LPMO	Endo-glucanase	Acetylxylan esterase	1,4-β-cellobiosidase	1,4-β-cellobiosidase	β-glucosidase
EC #	3.2.1.8	-	3.2.1.4	3.1.1.72	3.2.1.91	3.2.1.176	3.2.1.21
Abundance %	17.9	17.8	10.7	7.2	6.7	6.8	3.1

Scoring highest, endo 1,4 β-xylanase (EC 3.2.1.8); scoring second highest LPMO AA9; scoring third, endo 1,4 β glucanase (EC 3.2.1.4). The relative abundance values shown represent abundancies of enzyme proteins found to be secreted in the *R. rosea* experimental studies.

cellulolytic and xylanolytic enzymes but in pectinolytic enzymes. The three specialized parasitic species, *B. dendrobatidis*, *C. anguillulae*, and *R. allomycis*, had only few carbohydrate active enzymes in their genomes (57, 35 and 27, respectively; Supplementary Tab. A.8).

Overview of early lineage fungi cellulolytic, xylanolytic and pectinolytic enzyme capacity is shown in Tables 4a,b and 5a,b; (for detailed description, see Supplementary Text A.2 and Tab. A.8): β-glucan EC 3.2.1.6 active enzymes were predicted to be found only in *R. rosea* and *S. punctatus* (GH16). Enzymes with EC 3.2.1.39 function were found in *R. rosea* (GH16, GH55 and GH128) but in *G. prolifera* were represented by GH16 and GH128 only. The GH128, EC 3.2.1.6 was also found represented in genomes of *B. dendrobatidis*, *R. allomycis* and all three rumen fungi. Chitin: Four different enzyme functions, all with activities on chitin, were found. The richest portfolio of chitin-degrading and modifying enzymes were found in *B. dendrobatidis*: EC 3.2.1.14, 3.5.1.41, 3.2.1.52, 3.2.1.132 and 3.2.1.165, represented by GH18, CE4, GH20, GH46 and GH9, respectively. *Rhizophlyctis rosea*, *H. polyrhiza* and the three rumen fungi had only the first three of these; *B. dendrobatidis* only the first two. Both species of Blastocladiomycota also had these first two functions but also a GH2 of the 3.2.1.52 function. *Gonapodaya prolifera* had only one chitinase (viz. EC 3.2.1.14), but this function was represented by two different types of protein families, GH18 and GH19. The Cryptomycota, *R. allomycis* was found only to have one GH18 chitinase (EC 3.2.1.14). The defense related enzyme, lysozyme (EC 3.2.1.17) was found in *R. globosum* and *R. allomycis* (both belonging to GH19). *Neocallimastix californiae* also had a lysozyme but in this anaerobic rumen fungus it was a lysozyme belonging to GH73.

Unique CAZyme genes within the secretome of the five fungal phyla were primarily found in Neocallimastigomycota, *R. rosea* and *G. prolifera* (Chytridiomycota). As can be seen (Tables 4a,b and 5a,b), conserved patterns of similar enzyme protein family for similar enzyme function were found among the aerobic Chytridiomycota, Blastocladiomycota and Cryptomycota. The anaerobic rumen fungi had many unique

Table 4a – Overview of HotPep/dbCAN predicted plant cell wall degrading enzymes of zoosporic fungi. Species of Chytridiomycota.

		Chytridiomycota					
		Rhizophlyctis rosea	*Spizellomyces punctatus*	*Rhizoclosmatium globosum*	*Homoloaphlyctis polyrhiza*	*Batrachochytrium dendrobatidis*	*Gonapodya prolifera*
EC #	Function description	Enzyme families	Enzyme families	Enzyme families	Enzyme families	Enzyme families	Enzyme families
Cellulose							
3.2.1.4	Endo-1,4-β-D-glucanase	13 2 GH9 2 GH7 1 GH45 8 GH5	1 1 GH9				7 1 GH9 3 GH45 3 GH26
3.2.1.21	β-glucosidase	7 4 GH3 3 GH1	1 1 GH3	6 3 GH3 3 GH1		1 1 GH3	3 3 GH3
3.2.1.91	cellulose_1,4-β-cellobiosidase(non-red_end)	6 6 GH6					
3.2.1.176	cellulose_1,4-β-cellobiosidase_(red_end)	3 3 GH7					
AA9	LPMO	24 24 AA9					2 2 AA9
1.1.99.18	cellobiose_dehydrogenase_(acceptor)	1 1 AA8		1 1 AA3			2 2 AA3
1.1.99.29	pyranose_dehydrogenase_(acceptor)	1 1 AA3					
Hemi-cellulose							
3.1.1.6	acetylesterase	4 2 CE16 2 CE1					
3.1.1.72	Acetylxylan_esterase	6 2 CE1 4 CE5					3 3 CE4
3.1.1.73	feruloyl_esterase	1 1 GH10					
3.2.1.8	endo-1,4-β-xylanase	16 6 GH11 10 GH10					1 1 GH10
3.2.1.37	xylan_1,4-β-xylosidase	4 1 GH5 1 GH3 2 GH43					
3.2.1.55	α-N-arabinofuranosidase	1 1 GH43					
3.2.1.131	xylan_α-1,2-glucuronosidase	1 1 GH115					
3.2.1.151	xyloglucan-specific_endo-β-1,4-glucanase	3 3 GH74					
Lignin							
1.10.3.2	laccase		1 1 AA1	2 2 AA1		1 1 AA1	2 2 AA1
1.11.1.13	manganese_peroxidase			1 1 AA2			
CE15	4-O-methyl-glucuronoyl methylesterase	2 2 CE15					
Pectin							
3.1.1.11	pectinesterase	1 1 CE8					2 2 CE8
3.2.1.15	polygalacturonase	2 2 GH28		1 1 GH28			6 6 GH28
3.2.1.22	α-galactosidase			2 2 GH27	1 1 GH36		
3.2.1.23	β-galactosidase	1 1 GH35		4 4 GH35	1 1 GH35		1 1 GH2
3.2.1.40	α-L-rhamnosidase						
3.2.1.67	galacturan_1,4-α-galacturonidase	1 1 GH28					4 4 GH28
3.2.1.89	arabinogalactan_endo-β-1,4-galactanase	2 2 GH53					1 1 GH53
4.2.2.2	pectate_lyase	6 1 PL1 1 PL10 4 PL3					
4.2.2.10	pectin lyase						4 4 PL1
4.2.2.-	rhamnogalacturonan lyase	2 2 PL4					

enzymes: GH 36 α-galactosidase (EC 3.2.1.22); GH 78 α-L-rhamnosidase (EC 3.2.1.40); GH48 1,4 β-cellobiohydrolase, reducing-end active; GH26 1, 4 β-endoglucanase (EC 3.2.1.4); CE6 acetylxylan esterase (EC 3.1.1.72); GH8 endo 1,4 β-xylosidase (EC 3.2.1.8); GH32 β-fructofuranosidase (EC 3.2.1.26); and GH8 chitosanase (EC 3.2.1.132). Interestingly, monoblepharidomycete species, *G. prolifera*, also had unique enzymes, namely GH28 with galacturan 1,4 α-galacturonidase activity (EC 3.2.1.67) and GH62 α- N-arabinofuranosidase activity (EC 3.2.1.55), notably the latter not found in any other species studied.

A *PCA analysis* was made of all carbohydrate active enzyme hits. As shown in Supplementary Figs A.3-4, the *R. rosea* secretome occupied a truly outlier position as did the secretome of *N. californiae*, though to a lesser degree. The very significant composition of these two enzymatically strongest biomass degraders dominated the PCA plot to the extent that the relatedness and clustering of the groupings of the 12 species remained unresolved by PCA.

3.3. *LPMO AA9 and AA11 are found also in early lineage fungi*

We report here for the first time that lytic polysaccharide monooxygenases, LPMOs, are present in early lineage fungi Chytridiomycota (AA9 in *R. rosea* and *G. prolifera*), and that AA11 is found in Cryptomycota (*R. allomycis*). *Rhizophlyctis rosea* has a high number (24) of AA9 in its genome; and *G. prolifera* was found to have two AA9 (see Table 4a,b). Further, wild type, homolog expression and secretion of AA9 proteins were confirmed by MS for *R. rosea*, thus experimentally demonstrating the presence of LPMO proteins in *R. rosea* culture broth. Notably, AA9 was documented by MS to be among the most abundant proteins in the secretome of *R. rosea* (Tables 2 and 3). These first records of LPMOs in early lineage fungi were made by using the bioinformatics approach described by Busk and Lange, (2015).

Table 6 gives an overview of the more detailed information about the LPMO, AA9 and AA11 found in the zoosporic early lineage fungi *R. rosea*, *G. prolifera* and *R. allomycis*. The functional overview of LPMO diversity was adapted from Busk and Lange (2015). As Table 6 shows, the LPMO genes found represent a rich diversity, illustrated by their distribution over a wide spectrum of different LPMO PPR groups, including both C1 and C4 oxidizing types. Further, the variation in CBM1 domain and signal peptide occurrences is also apparent from Table 6. No LPMO were found in the genomes of any of the three rumen fungi studied; and no LPMO AA11 was found in the skin-invading frog pathogen, *B. dendrobatidis*.

Fig. 3 presents a (proportional) radial phylogram tree of fungal AA9. The tree includes a broad selection of fungal, ascomycetous and basidiomycetous AA9 sequences, along with the 20 (MS confirmed) new AA9 LPMO sequences from *R. rosea*. The tree consists of two all-fungal clades (i.e. including all three Asco-, Basidio- and Chrytridiomycota AA9); one clade with Asco- and Basidiomycota AA9, but no *R. rosea* AA9; a cluster of highly diverse, unique, early branching Chytridiomycota AA9 can also be seen (to the right). The resolution of this part of the tree is improved in the presentation of the same data given in Supplementary Fig. A.5 (here including boot-strap values).

Table 4b – Overview of HotPep/dbCAN predicted plant cell wall degrading enzymes of zoosporic fungi. Species of Neocallimastigiomycota, Blastocladiomycota and Cryptomycota.

		Neocallimastigomycota			Blastocladiomycota		Cryptomycota
		Piromyces finnis	*Anaeromyces robustus*	*Neocallimastix californiae*	*Allomyces macrogynus*	*Catenaria anguillulae*	*Rozella allomycis*
EC #	**Function description**	**Enzyme families**	**Enzyme families**	**Enzyme families**	**Enzyme families**	**Enzyme families**	**Enzyme families**
Cellulose							
3.2.1.4	Endo-1,4-β-D-glucanase	45 11 GH9 12 GH45 22 GH5	40 9 GH9 11 GH45 20 GH5	73 12 GH9 20 GH45 1 GH26 40 GH5			
3.2.1.21	β-glucosidase	24 14 GH3 10 GH1	20 14 GH3 6 GH1	58 42 GH3 16 GH1			
3.2.1.91	cellulose1,4-β-cellobiosidase(non-red_end)	20 20 GH6	11 11 GH6	25 25 GH6			
3.2.1.176	cellulose_1,4-β-cellobiosidase_(red_end)	12 **12 GH48**	6 **6 GH48**	20 **20 GH48**			
AA11	LPMO						2 **2 AA11**
1.1.99.18	cellobiose_dehydrogenase_(acceptor)						
1.1.99.29	pyranose_dehydrogenase_(acceptor)						
Hemi-cellulose							
3.1.1.6	acetylesterase			2 2 CE16	1 1 CE16	1 1 CE16	
3.1.1.72	Acetylxylan_esterase	12 1 CE2 2 CE4 9 CE6	34 11 CE2 12 CE4 11 CE6	41 5 CE2 22 CE4 14 CE6			
3.1.1.73	feruloyl_esterase	4 4 CE1	11 11 CE1	6 **6 GH10**			
3.2.1.8	endo-1,4-β-xylanase	55 36 GH11 18 GH1 1 GH8	45 29 GH11 14 GH1 2 GH8	70 22 GH11 1 GH3 1 GH8 46 GH10	4 4 GH11		2 **2 CE4**
3.2.1.37	xylan_1,4-β-xylosidase	2 2 GH43	7 1 GH120 6 GH43	11 2 GH3 9 GH43			
3.2.1.55	α-N-arabinofuranosidase	8 8 GH43	4 4 GH43	13 13 GH43			
3.2.1.131	xylan_α-1,2-glucuronosidase	2 2 GH115	1 1 GH115	7 7 GH115			
3.2.1.151	xyloglucan-specific_endo-β-1,4-glucanase	1 1 GH74	1 1 GH74	4 4 GH74			
Lignin							
1.10.3.2	laccase						
1.11.1.13	manganese_peroxidase						
CE15	4-O-methyl-glucuronoyl methylesterase	2 2CE15	2 2CE15	6 6CE15			
Pectin							
3.1.1.11	pectinesterase						
3.2.1.15	polygalacturonase			1 1 GH28	3 3 GH28		
3.2.1.22	α-galactosidase						
3.2.1.23	β-galactosidase	1 1 GH2	1 1 GH2	7 7 GH2			
3.2.1.40	α-L-rhamnosidase		1 1 GH78	1 1 GH78			
3.2.1.67	galacturan_1,4-α-galacturonidase						
3.2.1.89	arabinogalactan_endo-β-1,4-galactanase	1 1 GH53	1 1 GH53	3 3 GH53			
4.2.2.2	pectate_lyase	7 4 PL1 3 PL3	4 3 PL1 1 PL3	25 19 PL1 6 PL3		1 1 PL3	
4.2.2.10	pectin lyase				2 2 PL1		
4.2.2.-	rhamnogalacturonan lyase						

Table 4a,b include only a selection of the annotated CAZy enzyme hits, annotated to functions. A more comprehensive list is given in Table S5. Color code of fonts: red, unique enzyme, only found in one species; green, only found in one species of the phylum but also in other phyla; blue, only found in one phylum; black, found in all species studied.

Table 5a – Overview of HotPep/dbCAN discoveries of enzymes degrading storage materials and modifying fungal cell wall components. Chytridiomycota species.

		Chytridiomycota					
		Rhizophlyctis rosea	*Spizellomyces punctatus*	*Rhizoclosmatium globosum*	*Homoloaphlyctis polyrhiza*	*Batrachochytrium dendrobatidis*	*Gonapodya prolifera*
EC #	**Function description**	**Enzyme families**	**Enzyme families**	**Enzyme families**	**Enzyme families**	**Enzyme families**	**Enzyme families**
Starch							
3.2.1.1	α-amylase	1 1 GH13	2 2 GH13				
3.2.1.3	glucan_1,4-α-glucosidase	1 1 GH15	1 1 GH15	4 4 GH15	1 1 GH15	1 1 GH15	1 1 GH15
3.2.1.20	α-glucosidase	2 **1 GH31** 1 GH13	3 **3 GH31**	1 **1 GH31**	1 **1 GH31**	1 **1 GH31**	1 **1 GH31**
3.2.1.28	α,α-trehalase	2 **2 GH37**	3 **3 GH37**	1 **1 GH37**	1 **1 GH37**	1 **1 GH37**	2 **2 GH37**
Mannan							
3.2.1.24	α-mannosidase	2 2 GH38	1 1 GH38	3 3 GH38	1 1 GH38		1 1 GH38
3.2.1.25	β-mannosidase	2 2 GH2	1 1 GH2	1 1 GH2	1 1 GH2		
3.2.1.78	mannan_endo-1,4-β-mannosidase	6 4 GH5 2 GH26	1 1 GH5	1 1 GH5			
3.2.1.113	mannosyl-oligosaccharide_1,2-α-mannosidase	3 **3 GH47**	4 **4 GH47**	8 **8 GH47**	6 **6 GH47**	5 **4 GH47** 1 GH92	5 **5 GH47**
β-glucan							
3.2.1.6	endo-1,3(4)-β-glucanase - laminarinase	1 1 GH16	1 1 GH16				
3.2.1.39	glucan_endo-1,3-β-D-glucosidase	3 1 GH128 1 GH16 **1 GH55**		1 1 GH16			3 2 GH128 1 GH16
Chitin							
3.2.1.14	chitinase	7 **7 GH18**	2 **2 GH18**	22 **22 GH18**	2 **2 GH18**	3 **3 GH18**	13 **12 GH18** **1 GH19**
3.5.1.41	chitin_deacetylase	5 5 CE4	6 6 CE4	1 1 CE4	8 8 CE4	10 10 CE4	
3.2.1.52	β-N-acetylhexosaminidase	1 1 GH20	1 1 GH20	11 11 GH20	1 1 GH20		
3.2.1.132	chitosanase		1 1 GH8	1 1 GH46			
3.2.1.165	exo-1,4-β-D-glucosaminidase			1 **1 GH9**			

3.4. *Constructing phylogenetic trees for selected enzyme proteins of early lineage fungi*

Phylogenetic trees of homolog proteins from all Biological Kingdoms are shown in Figs 3, 4 and 6; and in Supplementary Figs A.5, A.6. The trees visualize the global diversity and relatedness of the enzymes from the zoosporic early lineage fungi to enzymes derived from other types of organisms (plants, animals, higher fungi, bacteria, and archaea).

In the phylogenetic tree of GH5 (EC3.2.1.4) presented in Fig. 4, sequences of *R. rosea* are embedded in fungal clades of the upper half of the tree, which also contain ascomycetous,

Table 5b – Overview of HotPep/dbCAN discoveries of enzymes degrading storage materials and modifying fungal cell wall components. Neocallimastigomycota, Blastocladiomycota and Cryptomycota species.

		Neocallimastigomycota			Blastocladiomycota		Cryptomycota
		Piromyces finnis	*Anaeromyces robustus*	*Neocallimastix californiae*	*Allomyces macrogynus*	*Catenaria anguillulae*	*Rozella allomycis*
EC #	**Function description**	**Enzyme families**	**Enzyme families**	**Enzyme families**	**Enzyme families**	**Enzyme families**	**Enzyme families**
Starch							
3.2.1.1	α-amylase		1 1 GH13			3 3 GH13	
3.2.1.3	glucan_1,4-α-glucosidase				8 8 GH15	2 2 GH15	
3.2.1.20	α-glucosidase	3 **2 GH31** 1 GH13	5 **5 GH31**	10 **9 GH31** 1 GH13	3 **3 GH31**	1 **1 GH31**	2 **2 GH31**
3.2.1.28	α,α-trehalase	1 **1 GH37**	1 **1 GH37**	1 **1 GH37**	2 **2 GH37**	1 **1 GH37**	2 **2 GH37**
Mannan							
3.2.1.24	α-mannosidase	1 1 GH38	2 2 GH38	1 1 GH38	5 5 GH38		1 1 GH38
3.2.1.25	β-mannosidase						
3.2.1.78	mannan_endo-1,4-β-mannosidase	4 1 GH5 3 GH26	6 4 GH5 2 GH26	14 7 GH5 7 GH26	2 2 GH26		
3.2.1.113	mannosyl-oligosaccharide_1,2-α-mannosidase	1 **1 GH47**	4 **4 GH47**	6 **6 GH47**	2 **2 GH47**	2 **2 GH47**	3 **3 GH47**
β-glucan							
3.2.1.6	endo-1,3(4)-β-glucanase - laminarinase						
3.2.1.39	glucan_endo-1,3-β-D-glucosidase	3 3 GH16	3 3 GH16	4 4 GH16			
Chitin							
3.2.1.14	chitinase	6 **6 GH18**	8 **8 GH18**	16 **16 GH18**	6 **6 GH18**	6 **6 GH18**	2 **2 GH18**
3.5.1.41	chitin_deacetylase	6 6 CE4	3 3 CE4	18 18 CE4	5 5 CE4	4 4 CE4	
3.2.1.52	β-N-acetylhexosaminidase				2 2 GH20	1 1 GH20	
3.2.1.132	chitosanase			1 1 GH8			
3.2.1.165	exo-1,4-β-D-glucosaminidase						

Table 5a,b include only a selection of the annotated CAZy enzyme hits, annotated to functions. A more comprehensive list is given in Table S5. Color code of fonts: red, unique enzyme, only found in one species; green, only found in one species of the phylum but also in other phyla; blue, only found in one phylum; black, found in all species studied.

basidiomycetous and zygomycetous GH5. However, in contrast, the GH5 from the anaerobic species studied are found embedded in a clade dominated by bacterial proteins, originating predominantly from the cow rumen habitat. Besides these two major occurrences of early lineage fungal GH5, one copy of *R. rosea* GH5 is found in a predominantly bacterial GH5 clade together with one GH5 protein from a zygomycetous species (*Umbelopsis ramanniana*) and an Archaea (*Pyrococcus abyssi*) GH5 protein. No GH5 with predicted EC 3.2.1.4 endoglucanase activity was found in *G. prolifera* genome.

The GH5 phylogenetic tree in Fig. 4 and the sub-tree in Supplementary Fig. A.6, based on same data, both illustrate two possible horizontal transfer (HGT) events. The data provide strong evidence for HGT events between rumen bacteria (*Fibrobacter succinogenes*) and the rumen zoosporic fungus *O. joyonii*. There is likewise strong evidence of HGT between rumen bacterium *Fibrobacter succinogenes* and rumen protozoan *Epidinium ecaudatum*.

To investigate further the difference between the *R. rosea* GH5 copies found in the fungal clade and one of the sequences found embedded in the bacteria-dominated clade, a comparison was made of their gene structure (see Fig. 5). The bacterial related *R. rosea* GH5 gene had only one intron and no CBM. By contrast, the fungal related *R. rosea* GH5 had a CBM1 domain and up to many introns and additional gene domain inserts.

The phylogenetic tree of GH13 (EC 3.2.1.1), shown in Fig. 6 in two versions, illustrates the extremely broad diversity of enzymes derived from the different phyla of early lineage fungi: GH13 from early lineage fungi are found in most parts of the radial tree. GH13 of Neocallimastigomycota and Blastocladiomycota are embedded in two different bacterial clades; while GH13 of *R. rosea* are found on a separate, very basal part of an all-fungal clade. However, in the opposite part of the tree, embedded in a bacterial clade is another GH13 gene variant of *R. rosea*. Adjacent to this is GH13 of the Chytrid, *S. punctatus*. *G. prolifera* GH13 is found separately.

4. Discussion

The investigations reported here of secreted carbohydrate active enzymes (CAZy) from early lineage fungi provide strong evidence that fungal (Asco- and Basidiomycetes) plant cell wall degrading enzymes share a common ancestor with enzymes from aerobic zoosporic early lineage fungi. Phylogenetic analysis revealed that the major part of *R. rosea* and *G. prolifera* plant cell wall degrading enzymes (Chang *et al.*, 2015) is embedded in or placed in clades closely associated to the clades of homolog enzymes from crown fungi (i.e. ascomycetous and basidiomycetous cellulolytic, xylanolytic and pectinolytic enzymes). However, we have also shown that the CAZy secretomes have developed very differently in the four phyla. The phylogenetic studies presented here show that plant biomass degrading enzymes from anaerobic rumen fungi are embedded in or placed in clades close to rumen bacteria (data shown for GH5 and GH13). This support the notion of the anaerobic Neocallimastigomycota being only more distantly related to the rest of the early lineage fungi. The results here reported suggest the common ancestor of CAZy cell wall degrading enzymes of rumen fungi may most likely have been of bacterial origin, with the genes initially acquired through HGT in the aqueous rumen habitat. Such dynamic

Table 6 – Overview of early lineage LPMO genes. *R. rosea* has 20 different expressed AA9 proteins!

AA family	Organism	PPR group	Observed by MS	Signal peptide	CBM family	dbCAN domain	Genbank accession number
AA9	*Rhizophlyctis rosea*	3	+	+	–	AA9	MF432135
		3	+	+	–	AA9	MF432148
		3	+	+	–	AA9	MF432140
		3	+	+	–	AA9	MF432144
		4	+	+	CBM1	AA9	MF432137
		5	+	–	–	AA9	MF432139
		5	+	–	CBM1	AA9	MF432136
		5	+	+	CBM1	AA9	MF432138
		8	+	+	–	AA9	MF432146
		8	+	+	CBM1	AA9	MF432133
		10	+	+	–	AA9	MF432150
		11	+	–	–	AA9	MF432143
		29	+	+	–	AA9	MF432149
		34	+	+	CBM1	AA9	MF432145
		35	+	+	–	AA9	MF432134
		35	+	+	–	AA9	MF432141
		39	+	–	CBM1	AA9	MF432147
		39	+	+	CBM1	AA9	MF432131
		–	+	+	CBM1	AA9	MF432142
		–	+	+	–	AA9	MF432132
	Gonapodya prolifera	2	N/A	+	CBM1	AA9	KXS12868.1
		14	N/A	+	CBM1	AA9	KXS20077.1
AA11*	*Rozella allomyces*	2	N/A	+	–	AA11	EPZ34955.1
		4	N/A	–	–	AA11	EPZ33847.1
		–	N/A	–	–	AA11	EPZ31597.1
		–	N/A	–	CBM19	AA11	EPZ34988.1

*The diversity of R. rosea AA9 is illustrated by the fact that they belong to ten different PPR groups *sensu* Busk and Lange, 2015. *G. prolifera* has only two LPMO AA9s belonging to two different PPR LPMO AA9 groupings, both of which are not found represented in *R. rosea*. Both of the AA9s of *G. prolifera* have a CBM1, while only 8 of the 20 AA9s in *R. rosea* have CBM1. One LPMO/AA11 from *R. allomycis* (Cryptomycota) was observed to have a chitin-binding domain; none of the *R. allomycis* AA11 was found to have a CBM1 binding domain.

interaction between rumen bacteria and the zoosporic rumen fungi may be an ongoing process. HGT is assumed to function in parallel with the fungal evolutionary mechanism of gene copy/gene loss (Braun *et al.*, 2000; Wapinski *et al.*, 2007).

The current comparative study of the phylogenetic variation among the homolog variants of GH5 (EC 3.2.1.4) shows that most of *R. rosea* GH5 proteins are placed within all-fungal clades (Fig. 4). However, one of the GH5 copies is embedded in an all-bacterial clade, and the gene structure of this gene is closer to the structure of a bacterial than a fungal gene (Fig. 5). These observed gene structure differences between different copies of GH5 suggest HGT from bacteria. Apparently, it is also an evolutionary gain for an aerobic chytrid like *R. rosea* to acquire more copies of homolog genes even though several copies are already in the genome.

Comparison between the two primary carbohydrate metabolizing Chytridiomycota included in this study, the terrestrial *R. rosea* and the aquatic *G. prolifera*, shows that the full spectrum of cellulolytic and xylanolytic plant cell wall degrading enzymes was found only in *R. rosea*. Comparison between only one aquatic and one terrestrial species is however not sufficient for conclusions to be drawn about enzyme secretome evolution. More plant biomass metabolizing, zoosporic fungi, both terrestrial and aquatic (e.g. belonging to Cladochytriales (Mozley-Standridge *et al.*, 2009), must be genome sequenced and functionally analyzed before any hypothesis can be formulated. Such broader studies would be highly interesting as it could provide insight in how the key cellulolytic and xylanolytic plant cell wall degrading enzymes evolved: Whether such enzymes were developed and maintained through speciation; and/or acquired through HGT (from which types of organisms?) in the transition from aquatic to terrestrial plant biomass degrading lifestyle. By contrast, a significant part of the mannan and starch degrading enzymes, and fungal cell wall modifying chitinases, were shared across all four early lineage fungal phyla investigated in this study (Table 5a,b); notably for these types of enzymes the rumen fungi are clearly related to the other fungal phyla. The data support a similar interpretation of the occurrence of the fungal carbohydrate binding domain CBM1, which is known to be an add-on domain of many CAZy enzymes in higher fungi (http://www.cazy.org/, n. d.). CBM1 was also found to be present widely in enzymes of early lineage fungi in the current work, indicating a common ancestor to CBMs among all parts of the Fungal Kingdom.

HGT has recently been shown to play a more significant role in evolution of the Fungal Kingdom than previously anticipated (Howlett *et al.*, 2007; Rosewich and Kistler, 2000; Schmitt and Lumbsch, 2009; Slot and Hibbett, 2007; Xie *et al.*, 2008). The current study includes observations of specific HGT events (see Fig. 4) which together with several other observations of comparative genome sequence analysis (Tables 4 and 5) suggest that HGT plays a prominent role in the evolution of the metabolic enzyme secretome of early lineage fungi. The

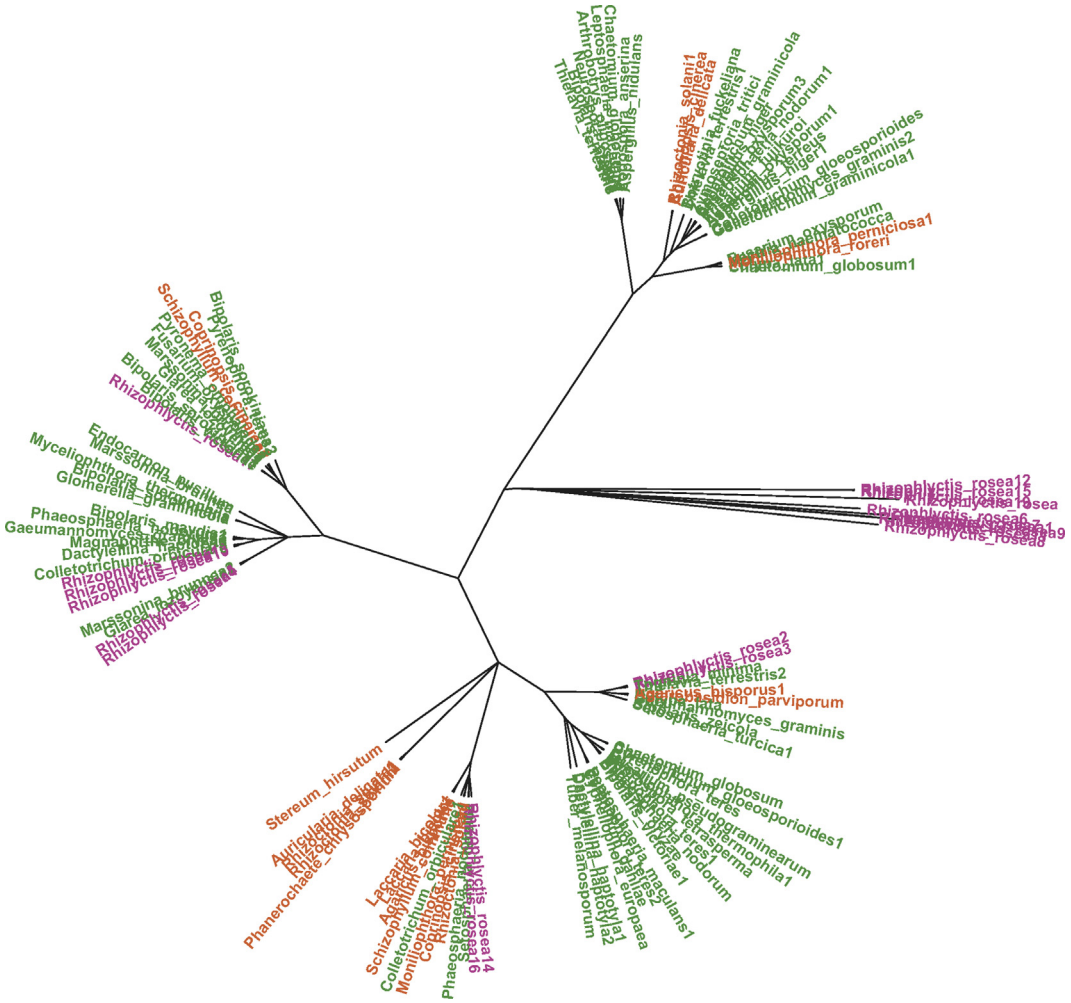

Fig. 3 – A radial phylogenetic tree of fungal AA9 LPMOs, including a broad selection of fungal, Ascomycota and Basidiomycota, AA9 LPMO sequences, along with the 20 new AA9 LPMO sequences found expressed in this study of wild type *R. rosea*. The tree has five major clades: three all-fungal clades (including asco-basidio- and Chytridiomycota AA9); one clade with AA9 from only Ascomycota and Basidiomycota. And pointing to the right, a cluster of seven, early branches of *R. rosea* AA9s; these seven closely clustered branches are unique for R. rosea and as appears only distantly related to the Ascomycota and Basidiomycota LPMO AA9s. Improved resolution of this part of the tree is given in Fig. S4, where also bootstrap values are included. Color code: green, Ascomycota AA9; orange, Basidiomycota AA9; purple, Chytridiomycota AA9.

phylogenetic trees of selected proteins provide basis for suggesting that HGT events have also taken place between other types of organisms; for example, HGT of GH5 (Fig. 4) from *Fibrobacter succinogenes* to the rumen fungus *O. joyonii* as well as to the dinoflagellate, *Epidinium*, and HGT of GH5 from rumen bacteria to the rumen fungi, *Neocallimastix*, *Orpinomyces* and *Piromyces*. Evolution of early lineage fungi from an aquatic environment to a terrestrial lifestyle and metabolizing lignocellulosic plant cell wall materials may have involved and required disruptive evolutionary mechanisms such as HGT.

Based on the above we conclude that occurrence of HGT is likely to be more prominent among aquatic organisms. Interestingly, a series of studies on the role of HGT in another group of aquatic organisms (the Oomycetes) has been published recently (Richards *et al.*, 2011; Savory *et al.*, 2015); such studies support our hypothesis that HGT occur more frequently among aquatic microbial life forms.

In the present study we document for the first time that oxidative LPMO enzymes are present in the early lineage fungi and not only in the higher Dikarya fungi, the ascomycetes and basidiomycetes (Beeson *et al.*, 2012; Busk and Lange, 2015). LPMO/AA9 was found in both the chytridiomycetes (*R. rosea*) and the monoblepharidiomycetes (*G. prolifera*) of the Chytridiomycota. Further, AA11 (Hemsworth *et al.*, 2014) was found in *R. allomycis*, positioned at the most basal part of the fungal evolutionary tree, the Cryptomycota. Interestingly, *R. rosea* was predicted to have as many as 24 different AA9 proteins in its genome. With this result we show that LPMOs in early lineage fungi also share the pattern, typically observed in higher fungi, that strong lignocellulosic biomass degraders in particular have several (or even many) monooxygenase genes (Busk and Lange, 2015). The sequence diversity of AA9 found in *R. rosea* is striking, illustrated by the fact that the 20 AA9 enzyme proteins found by MS analysis to be present in

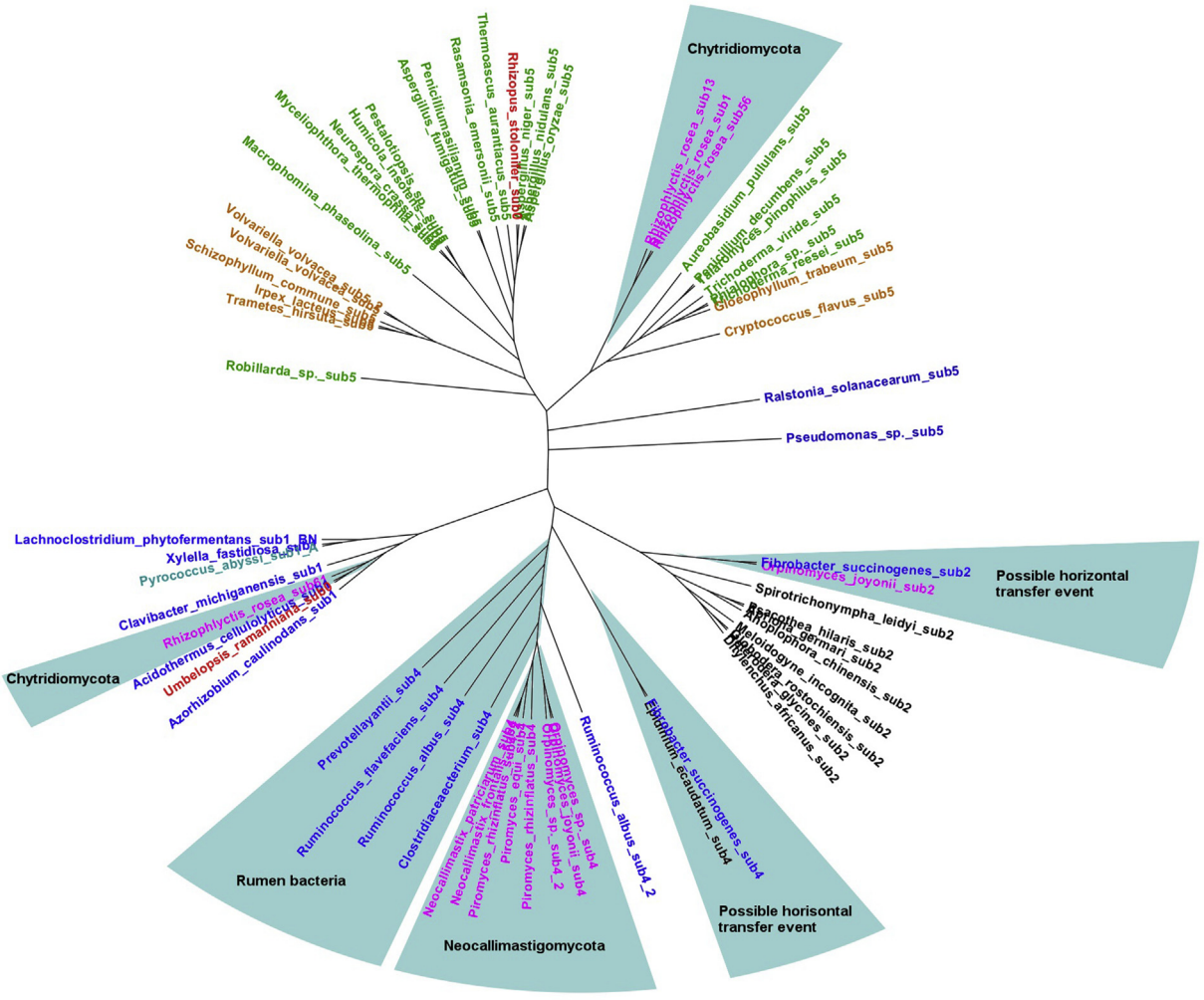

Fig. 4 – GH5 of early lineage fungi. The GH5 (EC 3.2.1.4) 1,4 β-endoglucanases of the chytrid *R. rosea* are embedded in an "all fungal" clade, containing also ascomycetous, basidiomycetous and zygomycetous GH5 endoglucanases. GH5 endoglucanases of anaerobic rumen fungi (Neocladiomastigomycota) are found in a clade dominated by rumen bacterial GH5 proteins. Besides these two major occurrences of early lineage fungal GH5, one copy of *R. rosea* GH5 is found in a predominantly bacterial GH5 clade (to the left) together with one GH5 protein from a zygomycetous species (*Umbelopsis ramanniana*) and an Archaea (*Pyrococcus abyssi*) GH5 protein. Two possible horizontal gene transfer events are indicated (between rumen bacteria and rumen fungi and between rumen bacteria and rumen protozoa). Color code: Purple, early lineage fungi; red, Zygomycota; green, Ascomycota. orange, Basidiomycota; blue, Bacteria; turquoise, Archaea; dark green, Planta; black, Animalia.

the secretome belong to 10 different AA9 PPR-groups. Notably, the two AA9 predicted proteins in *G. prolifera* belong to two different PPR-groups, neither of which were found in *R. rosea*; furthermore both of the two AA9 of *G. prolifera* have a carbohydrate binding domain, CBM1, while only 8 of the 20 AA9 in *R. rosea* have CBM1. Interestingly, none of the AA11 found in *R. allomycis* in this study had CBM domains, but one of the AA11 was predicted (Table 6) to have a chitin-binding domain, CBM19 (see also (Eijsink *et al.*, 2010; Várnai *et al.*, 2014)). In summary, we conclude that the phylogenetic studies of AA9 proteins of *R. rosea* and *G. prolifera* (Fig. 3) suggest a common ancestor with AA9 found in ascomycetes and basidiomycetes. A shared common ancestor also among all fungal AA11 is presented only as a hypothesis due to insufficient data. Notably, no AA11 LPMO was found in the genome of the frog pathogen, *B. dendrobatidis*, in contrast to fungal dermatophytes and *Onygena corvina* (Busk and Lange, 2015; Lange *et al.*, 2016).

The *R. rosea* experimental studies together with bioinformatics and MS analysis of protein abundance in the culture broth gave another surprising result. The LPMO enzyme proteins constitute a very significant proportion of the entire *R. rosea* enzyme secretome: LPMO AA9 proteins made up ≈ 18 % of all carbohydrate active enzyme protein found.

The discovery of LPMOs from two different phyla of early lineage fungi was made by the use of HotPep/dbCAN analysis. HotPep identification of conserved patterns of peptides enabled a range of new discoveries which again enabled formation of AA9 expanded family. Based on the expanded AA9 family, we also mined the genome of the protozoa

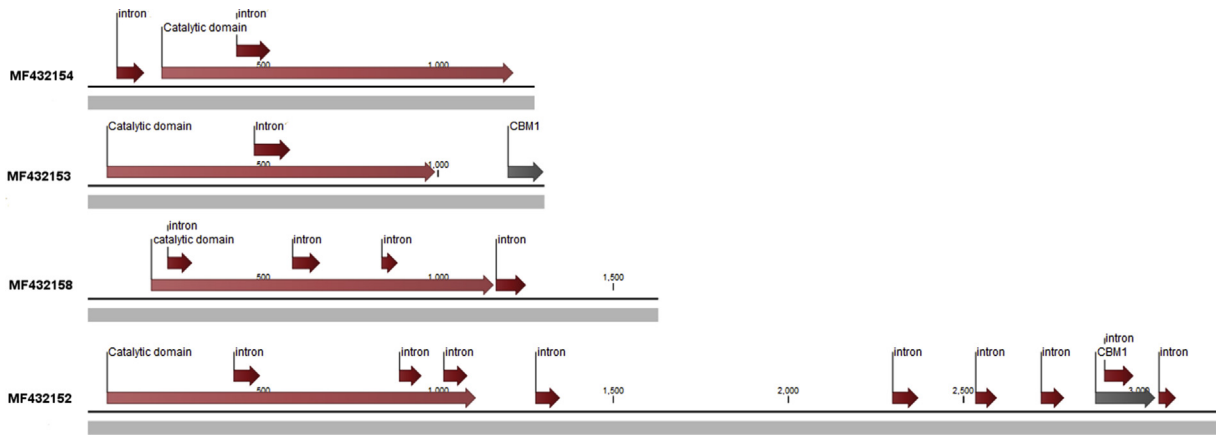

Fig. 5 – *R. rosea* GH5 gene structure: The bacterial-related GH5 variant of *R. rosea* (top gene, MF432154; placed in an all-bacterial clade in Fig. 4) has one intron in the catalytic domain and no CBM. The fungal related *R. rosea* GH5 variants have a CBM1 domain (MF432153, MF432152) and up to many introns and additional gene domain inserts (MF432153, MF432158, MF432152). This means that the gene structure of the bacterial-related *R. rosea* GH5 (Fig. 4) has a gene structure more closely related to bacterial genes, suggesting that it has been more recently acquired through horizontal gene transfer from bacteria.

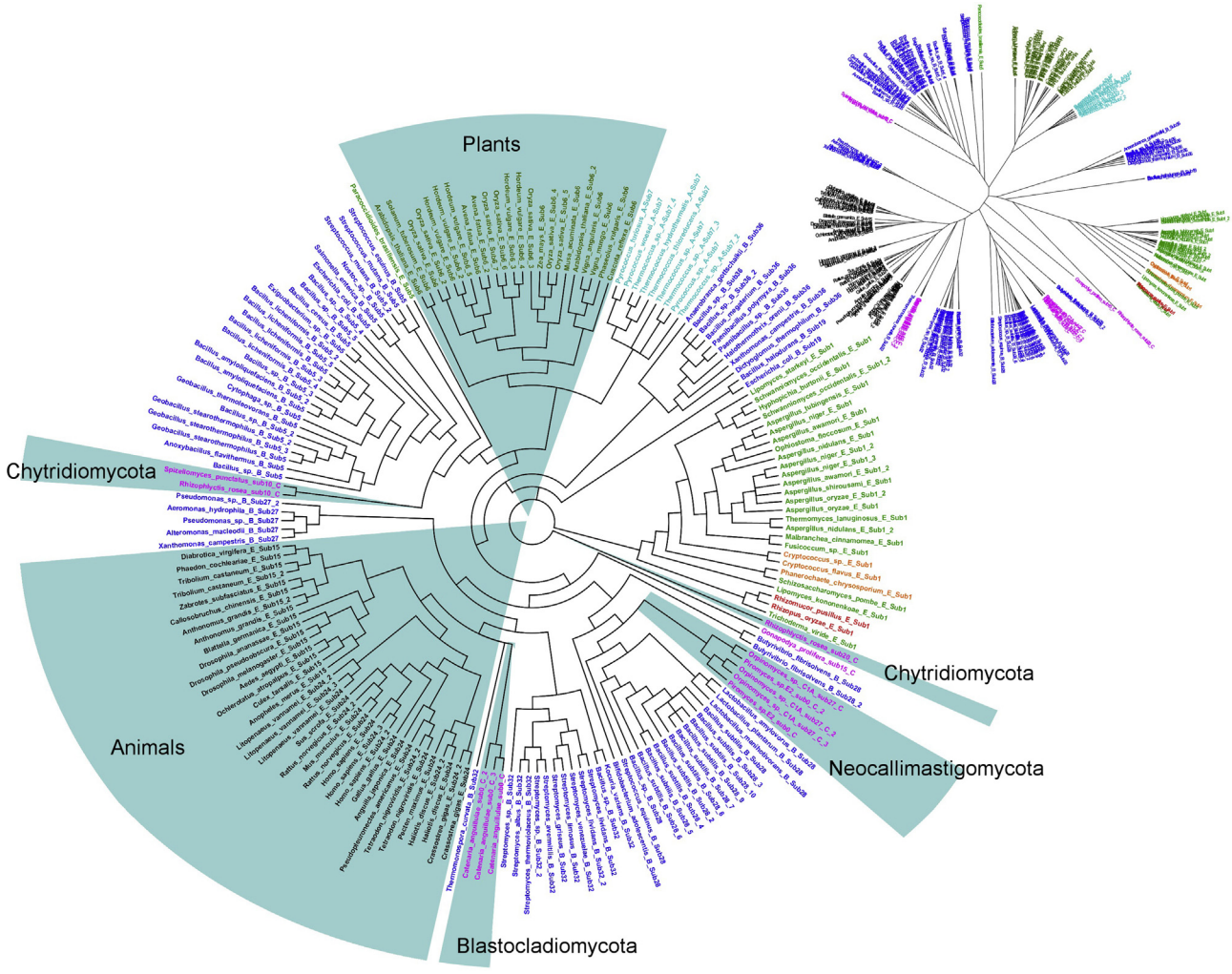

Fig. 6 – Phylogenetic tree of GH13, indicating the very broad diversity of enzymes of zoosporic fungi here found in all parts of the radial tree. GH13 of Neocallimastigomycota and Blastocladiomycota are both embedded in bacterial clade; GH13 of Chytridiomycota are placed in the all-fungal clade *and* in a bacterial clade. Color code: purple, early lineage fungi; red, Zygomycota; green, Ascomycota; orange, Basidiomycota; blue, bacteria; turquoise, Archaea; dark green, Planta; black, Animalia.

Dictyostelium discoideum (Eichinger *et al.*, 2005). Interestingly, for this soil-inhabiting amoeba, which is described as a rich cellulose degrader, we did not find any LPMOs. This result could support the interpretation that LPMOs have developed (or been acquired) at some time between when the protozoa lineage and the fungal lineage split off from the eukaryotes. Notably, LPMOs are found also in bacteria and viruses, and are described from a range of aquatic organisms (algae and animals) (Vaaje-Kolstad *et al.*, 2017). Both occurrences allow the possibility that LPMO can have been acquired or re-acquired by early lineage fungi through HGT, evolutionary paving the way for mobilizing cellulosic substrates.

Based on the findings that LPMOs are present in the early lineage fungi and that selected species of early lineage fungi have the fully fledged spectrum of cellulolytic, xylanolytic and pectinolytic cell wall degrading enzymes, we suggest the following evolutionary hypothesis: The challenge of decomposing the recalcitrant cellulose fibers of plant cell walls has been the driver for shaping the development of the entire spectrum of enzymes that degrade plant cell wall polymers. More specifically we suggest that the LPMOs could have played a key role in this evolution of microbial biomass conversion, essential for developing the biological Kingdoms. The radically different mechanism of LPMO enzymes creates synergy with the primarily hydrolytic cell wall degrading enzymes. Further, the many copies of LPMO could indicate that the LPMO mechanism is promiscuous, being active on many different types of plant (and animal) components going beyond cellulose and hemicellulose, possibly both by breaking bonds and by making room for the large enzyme proteins to establish contact with their substrates.

The outlier position of *R. rosea* is the most striking observation from the PCA plot (for strain description see (Gleason *et al.*, 2004). This position is underpinned by its full spectrum of cellulolytic, xylanolytic, and pectinolytic enzymes, as well as a host of proteins degrading mannan and starch. *R. rosea* in its evolutionary transition to the terrestrial life form (where it is a biomass degrader widely present in agricultural soils) has acquired or maintained a range of plant cell wall degrading enzymes. Notably, five of the nine different cellulolytic enzymes and 10 of the 13 xylanolytic enzymes of *R. rosea* are in this study unique for *R. rosea*; more specifically not found in the genomes of the other carbohydrate metabolizing species included in this study, neither in the rumen fungi nor in the *G. prolifera* genome. The other outlier in the PCA plots (Figs S3 and S4) is *N. californiae*, an anaerobic rumen fungus (Solomon *et al.*, 2016). *Neocallimastix californiae* (and other rumen fungi albeit to a lesser extent) has the most highly diversified spectrum of plant cell wall polymer degrading enzymes described, which is especially rich in cellulolytic and xylanolytic enzymes. *Gonapodaya prolifera* is an outstanding pectin-degrader (Chang *et al.*, 2015). Its spectrum of pectinolytic enzymes is, however, less diverse with regard to number of enzyme functions as compared to *R. rosea*. *Gonapodaya prolifera* appears to gain its pectinolytic strength from the many copies and functions of the GH28 enzyme. Its few but prominent cellulose and hemicellulose degrading enzymes suggest that it acts in synergy with other organisms (e.g. bacteria) in breaking down such plant polymers. *Synchytrium punctatus and R. globosum* appear from the secretome analysis to be specialized to degrade starch, and for the latter species also to degrade chitin. The specialized biotrophic and pathogenic species, *B. dendrobatidis*, *C. anguillulae*, *R. allomycis* and *H. polyrhiza*, have as expected from their lifestyle only few carbohydrate active enzymes, and basically no enzyme capacity for plant cell wall degradation (not relevant for its pathogenic specialization).

We document here that the enzymes of early lineage fungi represent a very rich pool of diversity of carbohydrate active enzymes of relevance also for industrial biomass conversion. A highly interesting aspect, possibly harboring significant additional enzyme diversity, is apparent from the enzyme variant/copy number heat map, Tables 4 and 5 This is however, only sporadically dealt with in the text above. From the phylogenetic trees, it appears that there is a high level of difference particularly between the *R. rosea* copies/variants of the same type of enzyme. This adds significant diversity to the enzyme discovery pool of early lineage fungi.

5. Conclusions and perspectives

The study of the enzymes of early lineage fungi, their diversity, function, phylogeny, evolution and applied potential is still only in its infancy. Seen from an evolutionary perspective aimed at contributing to resolving the lower lineages of the Fungal Tree of Life, the most interesting next steps would be to broaden such studies to include the enzyme profiles of Glomeromycota and Microsporidia. For elucidating the secretome evolution of zoosporic fungi, in the transition from aquatic to terrestrial life form the obvious next steps are to genome sequence more carbohydrate metabolizing aquatic and terrestrial zoosporic fungi, It further seems justified based on the current studies to suggest to include description of the enzyme activity profile to future descriptions of new fungal taxa as exemplified in the *Aspergillus hanckokii* paper (Pitt *et al.*, 2017). An LPMO-related puzzle worth investigating is that we have not found any enzyme mechanism to explain how rumen fungi under anaerobic conditions cope with breaking down recalcitrant cellulose from plant cell walls when they lack LPMOs; oxygenase activity does not normally function under anaerobic conditions (Vaaje-Kolstad *et al.*, 2017). A hypothesis could be that GH48 and/or rumen bacterial cellulosomes possess a stronger processivity hydrolysis mechanism.

Conflict of interest

No issues of conflicting interest have been declared or identified.

Role of authors

Lene Lange: Formulated the idea, conceptualized the study, supplied expert knowledge regarding laboratory handling of Early Lineage fungi and wrote the paper.

Bo Pilgaard: Carried out all the experimental work with *R. rosea*, cloning, expression and characterization of the proteins, including enzyme assaying; and was responsible for HotPep and dbCAN analysis of genomes of all of the 12 species studied.

Florian-Alexander Herbst: Responsible for the MS analysis of *R. rosea* secretome, including interpretation of the results and calculation of relative abundances.

Peter Kamp Busk: Developed —together with Lene Lange- the PPR-based sequence technology platform. More specifically developed the HotPep program used in the current paper, as basis for prediction of function directly from sequence.

Frank Gleason: Sampled, identified and described the strong biomass degrading strain of *R. rosea*; contributed information about ecology and lab-culturing of this zoosporic fungus.

Anders Gorm Pedersen: Constructed and interpreted the phylogenetic trees included in the study.

Acknowledgement

We thank Kristian Barrett, PhD student, for his kind assistance with the lay out of the phylogenetic trees; and Pia Nord-Larsen for technical assistance in preparing the manuscript. The study was supported by Innovation Fund Denmark, #0603-00522B (BIOVALUE SPIR).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fbr.2018.09.001.

REFERENCES

Andrews, S., 2010. Fast QC: a Quality-control Tool for High-throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Beeson, W.T., Phillips, C.M., Cate, J.H.D., Marletta, M.A., 2012. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. J. Am. Chem. Soc. 134. https://doi.org/10.1021/ja210657t.

Berbee, M.L., James, T.Y., Strullu-Derrien, C., 2017. Early Diverging Fungi: Diversity and Impact at the Dawn of Terrestrial Life. Annu. Rev. Microbiol.. https://doi.org/10.1146/annurev-micro-030117-020324.

Braun, E.L., Halpern, A.L., Nelson, M.A., Natvig, D.O., 2000. Large-scale comparison of fungal sequence information: Mechanisms of innovation in Neurospora crassa and gene loss in Saccharomyces cerevisiae. Genome Res. 10, 416—430. https://doi.org/10.1101/gr.10.4.416.

Busk, P.K., Lange, L., 2015. Classification of fungal and bacterial lytic polysaccharide monooxygenases. BMC Genomics 16, 368. https://doi.org/10.1186/s12864-015-1601-6.

Busk, P.K., Lange, L., 2013a. Cellulolytic potential of thermophilic species from four fungal orders. Amb. Express 3, 47. https://doi.org/10.1186/2191-0855-3-47.

Busk, P.K., Lange, L., 2013b. Function-based classification of carbohydrate-active enzymes by recognition of short, conserved peptide motifs. Appl. Environ. Microbiol. 79, 3380—3391. https://doi.org/10.1128/AEM.03803-12.

Busk, P.K., Lange, M., Pilgaard, B., Lange, L., 2014. Several genes encoding enzymes with the same activity are necessary for aerobic fungal degradation of cellulose in nature. PLoS One 9 (12), e11. https://doi.org/10.1371/journal.pone.0114138.

Busk, P.K., Pagès, M., 2002. In vivo footprinting of plant tissues. Plant Mol. Biol. Rep. 20, 287—297. https://doi.org/10.1007/BF02782464.

Busk, P.K., Pilgaard, B., Lezyk, M.J., Meyer, A.S., Lange, L., 2017. Homology to peptide pattern for annotation of carbohydrate-active enzymes and prediction of function. BMC Bioinf. 18, 214. https://doi.org/10.1186/s12859-017-1625-9.

Cantarel, B.L., Coutinho, P.M., 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. https://doi.org/10.1093/nar/gkn663.

Chang, Y., Wang, S., Sekimoto, S., Aerts, A.L., Choi, C., Clum, A., LaButti, K.M., Lindquist, E.A., Ngan, C.Y., Ohm, R.A., Salamov, A.A., Grigoriev, I.V., Spatafora, J.W., Berbee, M.L., 2015. Phylogenomic analyses indicate that early fungi evolved digesting cell walls of algal ancestors of land plants. Genome Biol. Evol. 7, 1590—1601. https://doi.org/10.1093/gbe/evv090.

Eddy, S.R., 2011. Accelerated profile HMM searches. PLoS Comput. Biol. 7. https://doi.org/10.1371/journal.pcbi.1002195.

Eichinger, L., Pachebat, J.A., Glöckner, G., Rajandream, M.-A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., Tunggal, B., Kummerfeld, S., Madera, M., Konfortov, B.A., Rivero, F., Bankier, A.T., Lehmann, R., Hamlin, N., Davies, R., Gaudet, P., Fey, P., Pilcher, K., Chen, G., Saunders, D., Sodergren, E., Davis, P., Kerhornou, A., Nie, X., Hall, N., Anjard, C., Hemphill, L., Bason, N., Farbrother, P., Desany, B., Just, E., Morio, T., Rost, R., Churcher, C., Cooper, J., Haydock, S., van Driessche, N., Cronin, A., Goodhead, I., Muzny, D., Mourier, T., Pain, A., Lu, M., Harper, D., Lindsay, R., Hauser, H., James, K., Quiles, M., Madan Babu, M., Saito, T., Buchrieser, C., Wardroper, A., Felder, M., Thangavelu, M., Johnson, D., Knights, A., Loulseged, H., Mungall, K., Oliver, K., Price, C., Quail, M.A., Urushihara, H., Hernandez, J., Rabbinowitsch, E., Steffen, D., Sanders, M., Ma, J., Kohara, Y., Sharp, S., Simmonds, M., Spiegler, S., Tivey, A., Sugano, S., White, B., Walker, D., Woodward, J., Winckler, T., Tanaka, Y., Shaulsky, G., Schleicher, M., Weinstock, G., Rosenthal, A., Cox, E.C., Chisholm, R.L., Gibbs, R., Loomis, W.F., Platzer, M., Kay, R.R., Williams, J., Dear, P.H., Noegel, A.A., Barrell, B., Kuspa, A., 2005. The genome of the social amoeba Dictyostelium discoideum. Nature 435, 43—57. https://doi.org/10.1038/nature03481.

Eijsink, V., Hoell, I., Vaaje-Kolstada, G., 2010. Structure and function of enzymes acting on chitin and chitosan. Biotechnol. Genet. Eng. Rev. 27, 331—366. https://doi.org/10.1080/02648725.2010.10648156.

Gleason, F.H., Letcher, P.M., McGee, P. a, 2004. Some Chytridiomycota in soil recover from drying and high temperatures. Mycol. Res. 108, 583—589. https://doi.org/10.1017/S0953756204009736.

Gleason, F.H., Scholz, B., Jephcott, T.G., van Ogtrop, F.F., Henderson, L., Lilje, O., Kittelmann, S., Macarthur, D.J., 2017. Key ecological roles for zoosporic true fungi in aquatic habitats. Microbiol. Spectr. 5.

Grigoriev, I.V., Nordberg, H., Shabalov, I., Aerts, A., Cantor, M., Goodstein, D., Kuo, A., Minovitsky, S., Nikitin, R., Ohm, R.A., Otillar, R., Poliakov, A., Ratnere, I., Riley, R., Smirnova, T., Rokhsar, D., Dubchak, I., 2012. The genome portal of the department of energy joint genome institute. Nucleic Acids Res. 40, D26—D32.

Hemsworth, G.R., Henrissat, B., Davies, G.J., Walton, P.H., 2014. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. Nat. Chem. Biol. 10, 122—126. https://doi.org/10.1038/nchembio.1417.

Hibbett, D.S., Binder, M., Bischoff, J.F., Blackwell, M., Cannon, P.F., Eriksson, O.E., Huhndorf, S., James, T.,

Kirk, P.M., Lücking, R., Thorsten Lumbsch, H., Lutzoni, F., Matheny, P.B., McLaughlin, D.J., Powell, M.J., Redhead, S., Schoch, C.L., Spatafora, J.W., Stalpers, J.A., Vilgalys, R., Aime, M.C., Aptroot, A., Bauer, R., Begerow, D., Benny, G.L., Castlebury, L.A., Crous, P.W., Dai, Y.-C., Gams, W., Geiser, D.M., Griffith, G.W., Gueidan, C., Hawksworth, D.L., Hestmark, G., Hosaka, K., Humber, R.A., Hyde, K.D., Ironside, J.E., Kõljalg, U., Kurtzman, C.P., Larsson, K.-H., Lichtwardt, R., Longcore, J., Miądlikowska, J., Miller, A., Moncalvo, J.-M., Mozley-Standridge, S., Oberwinkler, F., Parmasto, E., Reeb, V., Rogers, J.D., Roux, C., Ryvarden, L., Sampaio, J.P., Schüßler, A., Sugiyama, J., Thorn, R.G., Tibell, L., Untereiner, W.A., Walker, C., Wang, Z., Weir, A., Weiss, M., White, M.M., Winka, K., Yao, Y.-J., Zhang, N., 2007. A higher-level phylogenetic classification of the Fungi. Mycol. Res. 111, 509–547. https://doi.org/10.1016/j.mycres.2007.03.004.

Howlett, B.J., Idnurm, A., Heitman, J., 2007. Fungal pathogenesis: Gene clusters unveiled as secrets within the Ustilago maydis code. Curr. Biol. 17, R87–R90. https://doi.org/10.1016/j.cub.2006.11.047.

Huang, L., Zhang, H., Wu, P., Entwistle, S., Li, X., Yohe, T., Yi, H., Yang, Z., Yin, Y., 2018. DbCAN-seq: A database of carbohydrate-active enzyme (CAZyme) sequence and annotation. Nucleic Acids Res. 46, D516–D521. https://doi.org/10.1093/nar/gkx894.

Huang, Y., Busk, P.K., Herbst, F.A., Lange, L., 2015. Genome and secretome analyses provide insights into keratin decomposition by novel proteases from the non-pathogenic fungus Onygena corvina. Appl. Microbiol. Biotechnol. 99, 9635–9649. https://doi.org/10.1007/s00253-015-6805-9.

James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., Lumbsch, H.T., Rauhut, A., Reeb, V., Arnold, A.E., Amtoft, A., Stajich, J.E., Hosaka, K., Sung, G.-H., Johnson, D., O'Rourke, B., Crockett, M., Binder, M., Curtis, J.M., Slot, J.C., Wang, Z., Wilson, A.W., Schuszler, A., Longcore, J.E., O'Donnell, K., Mozley-Standridge, S., Porter, D., Letcher, P.M., Powell, M.J., Taylor, J.W., White, M.M., Griffith, G.W., Davies, D.R., Humber, R.A., Morton, J.B., Sugiyama, J., Rossman, A.Y., Rogers, J.D., Pfister, D.H., Hewitt, D., Hansen, K., Hambleton, S., Shoemaker, R.A., Kohlmeyer, J., Volkmann-Kohlmeyer, B., Spotts, R.A., Serdani, M., Crous, P.W., Hughes, K.W., Matsuura, K., Langer, E., Langer, G., Untereiner, W.A., Lucking, R., Budel, B., Geiser, D.M., Aptroot, A., Diederich, P., Schmitt, I., Schultz, M., Yahr, R., Hibbett, D.S., Lutzoni, F., McLaughlin, D.J., Spatafora, J.W., Vilgalys, R., 2006a. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443, 818–822.

James, T.Y., Letcher, P.M., Longcore, J.E., Mozley-Standridge, S.E., Porter, D., Powell, M.J., Griffith, G.W., Vilgalys, R., 2006b. A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). Mycologia 98, 860–871. https://doi.org/10.3852/mycologia.98.6.860.

James, T.Y., Pelin, A., Bonen, L., Ahrendt, S., Sain, D., Corradi, N., Stajich, J.E., 2013. Shared Signatures of Parasitism and Phylogenomics Unite Cryptomycota and Microsporidia. Curr. Biol. 23, 1548–1553. https://doi.org/10.1016/j.cub.2013.06.057.

Jephcott, T.G., Sime-Ngando, T., Gleason, F.H., Macarthur, D.J., 2016. Host–parasite interactions in food webs: Diversity, stability, and coevolution. Food Webs 6, 1–8. https://doi.org/10.1016/j.fooweb.2015.12.001.

Jones, M.D.M., Richards, T.A., Hawksworth, D.L., Bass, D., 2011. Validation and justification of the phylum name *Cryptomycota phyl.* nov. IMA Fungus Glob. Mycol. J. 2, 173–175. https://doi.org/10.5598/imafungus.2011.02.02.08.

Joneson, S., Stajich, J.E., Shiu, S.-H., Rosenblum, E.B., 2011. Genomic Transition to Pathogenicity in Chytrid Fungi. PLoS Pathog. 7, e1002338.

Käll, L., Krogh, A., Sonnhammer, E.L.L., 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids Res. 35, W429–W432. https://doi.org/10.1093/nar/gkm256.

Kämper, J., Kahmann, R., Bolker, M., Ma, L.-J., Brefort, T., Saville, B.J., Banuett, F., Kronstad, J.W., Gold, S.E., Muller, O., Perlin, M.H., Wosten, H.A.B., de Vries, R., Ruiz-Herrera, J., Reynaga-Pena, C.G., Snetselaar, K., McCann, M., Perez-Martin, J., Feldbrugge, M., Basse, C.W., Steinberg, G., Ibeas, J.I., Holloman, W., Guzman, P., Farman, M., Stajich, J.E., Sentandreu, R., Gonzalez-Prieto, J.M., Kennell, J.C., Molina, L., Schirawski, J., Mendoza-Mendoza, A., Greilinger, D., Munch, K., Rossel, N., Scherer, M., Vranes, M., Ladendorf, O., Vincon, V., Fuchs, U., Sandrock, B., Meng, S., Ho, E.C.H., Cahill, M.J., Boyce, K.J., Klose, J., Klosterman, S.J., Deelstra, H.J., Ortiz-Castellanos, L., Li, W., Sanchez-Alonso, P., Schreier, P.H., Hauser-Hahn, I., Vaupel, M., Koopmann, E., Friedrich, G., Voss, H., Schluter, T., Margolis, J., Platt, D., Swimmer, C., Gnirke, A., Chen, F., Vysotskaia, V., Mannhaupt, G., Guldener, U., Munsterkotter, M., Haase, D., Oesterheld, M., Mewes, H.-W., Mauceli, E.W., DeCaprio, D., Wade, C.M., Butler, J., Young, S., Jaffe, D.B., Calvo, S., Nusbaum, C., Galagan, J., Birren, B.W., 2006. Insights from the genome of the biotrophic fungal plant pathogen Ustilago maydis. Nature 444, 97–101.

Kassambara, A., Mundt, F., 2017. Package "factoextra." R Top. Doc..

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. 9, 286–298. https://doi.org/10.1093/bib/bbn013.

Kelley, D.R., Schatz, M.C., Salzberg, S.L., 2010. Quake: quality-aware detection and correction of sequencing errors. Genome Biol. 11, R116. https://doi.org/10.1186/gb-2010-11-11-r116.

Krijger, J.J., Thon, M.R., Deising, H.B., Wirsel, S.G.R., 2014. Compositions of fungal secretomes indicate a greater impact of phylogenetic history than lifestyle adaptation. BMC Genomics 15. https://doi.org/10.1186/1471-2164-15-722.

Lange, L., Huang, Y., Busk, P.K., 2016. Microbial decomposition of keratin in nature—a new hypothesis of industrial relevance. Appl. Microbiol. Biotechnol. 100, 2083–2096. https://doi.org/10.1007/s00253-015-7262-1.

Letcher, P.M., Powell, M.J., Barr, D.J.S., Churchill, P.F., Wakefield, W.S., Picard, K.T., 2008. Rhizophlyctidales—a new order in Chytridiomycota. Mycol. Res. 112, 1031–1048. https://doi.org/10.1016/j.mycres.2008.03.007.

Letcher, P.M., Powell, M.J., Churchill, P.F., Chambers, J.G., 2006. Ultrastructural and molecular phylogenetic delineation of a new order, the Rhizophydiales (Chytridiomycota). Mycol. Res. 110, 898–915. https://doi.org/10.1016/j.mycres.2006.06.011.

Lindgreen, S., 2012. AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC Res. Notes 5, 337. https://doi.org/10.1186/1756-0500-5-337.

Madsen, M.B., Birck, M.M., Fredholm, M., Cirera, S., 2009. Expression studies of the obesity candidate gene FTO in pig. Anim. Biotechnol. 21, 51–63. https://doi.org/10.1080/10495390903381792.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Bryant, S.H., 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43, D222–D226. https://doi.org/10.1093/nar/gku1221.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 17, 10–12. https://doi.org/10.14806/ej.17.1.200.

Mozley-Standridge, S.E., Letcher, P.M., Longcore, J.E., Porter, D., Simmons, D.R., 2009. Cladochytriales—a new order in Chytridiomycota. Mycol. Res. 113, 498–507. https://doi.org/10.1016/j.mycres.2008.12.004.

Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I.V., Dubchak, I., 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res. 42, D26–D31. https://doi.org/10.1093/nar/gkt1069.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Meth. 8, 785–786. https://doi.org/10.1038/nmeth.1701.

Pilgaard, B., 2014. Cloning, Expression and Characterization of a Glycoside Hydrolase Family 45 Enzyme from the Draft Genome Assembly of Rhizophlyctis Rosea. Aalborg University, Denmark.

Pitt, J.I., Lange, L., Lacey, A.E., Vuong, D., Midgley, D.J., Greenfield, P., Bradbury, M.I., Lacey, E., Busk, P.K., Pilgaard, B., Chooi, Y.-H., Piggott, A.M., 2017. Aspergillus hancockii sp. nov., a biosynthetically talented fungus endemic to southeastern Australian soils. PLoS One 12, e0170254.

Powell, M., Letcher, P., 2012. From zoospores to molecules. In: Systematics and Evolution of Fungi. Science Publishers, pp. 29–54. https://doi.org/10.1201/b11606-4.

Powell, M.J., Letcher, P.M., 2014. 6 Chytridiomycota, Monoblepharidomycota, and Neocallimastigomycota. In: McLaughlin, D.J., Spatafora, J.W. (Eds.), The Mycota. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 141–175. https://doi.org/10.1007/978-3-642-55318-9_6.

R core team, 2017. R: a Language and Environment for Statistical Computing. R Found. Stat. Comput. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Richards, T.A., Soanes, D.M., Jones, M.D.M., Vasieva, O., Leonard, G., Paszkiewicz, K., Foster, P.G., Hall, N., Talbot, N.J., 2011. Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. Proc. Natl. Acad. Sci. 108, 15258–15263. https://doi.org/10.1073/pnas.1105100108.

Rosewich, U.L., Kistler, H.C., 2000. Role of horizontal gene transfer in the evolution of fungi. Annu. Rev. Phytopathol. 38, 325–363. https://doi.org/10.1146/annurev.phyto.38.1.325.

RStudio Team, 2016. RStudio: Integrated Development for R [Online]. RStudio, Inc., Boston, MA. https://doi.org/10.1007/978-81-322-2340-5. http//www.rstudio.com RStudio,Inc., Boston, MA.

Ruggiero, M.A., Gordon, D.P., Orrell, T.M., Bailly, N., Bourgoin, T., Brusca, R.C., Cavalier-Smith, T., Guiry, M.D., Kirk, P.M., 2015. A higher level classification of all living organisms. PLoS One 10, e0119248.

Russ, C., Lang, B.F., Chen, Z., Gujja, S., Shea, T., Zeng, Q., Young, S., Cuomo, C.A., Nusbaum, C., 2016. Genome Sequence of Spizellomyces punctatus. Genome Announc. 4. https://doi.org/10.1128/genomeA.00849-16 e00849–16.

Savory, F., Leonard, G., Richards, T.A., 2015. The Role of Horizontal Gene Transfer in the Evolution of the Oomycetes. PLoS Pathog. 11, e1004805.

Schmitt, I., Lumbsch, H.T., 2009. Ancient Horizontal Gene Transfer from Bacteria Enhances Biosynthetic Capabilities of Fungi. PLoS One 4, e4437.

Slot, J.C., Hibbett, D.S., 2007. Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: A Phylogenetic Study. PLoS One 2, e1097.

Solomon, K.V., Haitjema, C.H., Henske, J.K., Gilmore, S.P., Borges-Rivera, D., Lipzen, A., Brewer, H.M., Purvine, S.O., Wright, A.T., Theodorou, M.K., Grigoriev, I.V., Regev, A., Thompson, D.A., O'Malley, M.A., 2016. Early-branching gut fungi possess large, comprehensive array of biomass-degrading enzymes. Science (80) 351, 1192–1195. https://doi.org/10.1126/science.aad1431.

Spatafora, J.W., Chang, Y., Benny, G.L., Lazarus, K., Smith, M.E., Berbee, M.L., Bonito, G., Corradi, N., Grigoriev, I., Gryganskyi, A., James, T.Y., O'Donnell, K., Roberson, R.W., Taylor, T.N., Uehling, J., Vilgalys, R., White, M.M., Stajich, J.E., 2016. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia 108, 1028–1046. https://doi.org/10.3852/16-042.

Stanke, M., Diekhans, M., Baertsch, R., Haussler, D., 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics 24, 637–644. https://doi.org/10.1093/bioinformatics/btn013.

Taylor, J.W., Berbee, M.L., 2006. Dating divergences in the Fungal Tree of Life: review and new analyses. Mycologia 98, 838–849.

Tyanova, S., Temu, T., Cox, J., 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat. Protoc. 11, 2301–2319. https://doi.org/10.1038/nprot.2016.136.

Vaaje-Kolstad, G., Forsberg, Z., Loose, J.S.M., Bissaro, B., Eijsink, V.G.H., 2017. Structural diversity of lytic polysaccharide monooxygenases. Curr. Opin. Struct. Biol. 44, 67–76. https://doi.org/10.1016/j.sbi.2016.12.012.

Várnai, A., Mäkelä, M.R., Djajadi, D.T., Rahikainen, J., Hatakka, A., Viikari, L., 2014. Carbohydrate-binding modules of fungal cellulases. occurrence in nature, function, and relevance in industrial biomass conversion. Adv. Appl. Microbiol. 88, 103–165. https://doi.org/10.1016/B978-0-12-800260-5.00004-8.

Wapinski, I., Pfeffer, A., Friedman, N., Regev, A., 2007. Natural history and evolutionary principles of gene duplication in fungi. Nature 449, 54–61. https://doi.org/10.1038/nature06107.

Xie, J., Fu, Y., Jiang, D., Li, G., Huang, J., Li, B., Hsiang, T., Peng, Y., 2008. Intergeneric transfer of ribosomal genes between two fungi. BMC Evol. Biol. 8, 87. https://doi.org/10.1186/1471-2148-8-87.

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., Xu, Y., 2012. DbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 40. https://doi.org/10.1093/nar/gks479.

Youssef, N.H., Couger, M.B., Struchtemeyer, C.G., Liggenstoffer, A.S., Prade, R.A., Najar, F.Z., Atiyeh, H.K., Wilkins, M.R., Elshahed, M.S., 2013. The genome of the anaerobic fungus Orpinomyces sp. strain C1A reveals the unique evolutionary history of a remarkable plant biomass degrader. Appl. Environ. Microbiol. 79, 4620–4634. https://doi.org/10.1128/AEM.00821-13.

Zerbino, D.R., Birney, E., 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829. https://doi.org/10.1101/gr.074492.107.

Lene Lange, Professor, PhD et Dr. scient., Center for BioProcess Engineering, Technical University of Denmark, DK2800 Lyngby, Denmark. **Lene Lange** has held full professorship positions at three Danish Universities, KU, AAU and DTU, and been Director of Research in both public and private research organizations. All through her career, she had her own experimental mycology research group, focusing on fungal interaction with their hosts and substrates. Current methodological approach: fast track discovery of new enzymes, predicting function directly from sequence; combined with mass spectrometry, phylogenetic analysis and activity screening for elucidating (meta-) secretome composition of taxonomic and ecological biomass conversion hotspots in Nature. Her interest in zoosporic fungi dates back to being research assistant for Professor F.K. Sparrow.