Metabolic potential, ecology and presence of endohyphal bacteria is reflected in genomic diversity of Mucoromycotina

Short title: Genomic diversity of Mucorales and Umbelopsidales

Anna Muszewska[1,*], Alicja Okrasińska[2], Kamil Steczkiewicz[1], Olga Drgas[1], Małgorzata Orłowska[1], Urszula Perlińska-Lenart[1],Tamara Aleksandrzak-Piekarczyk[1], Katarzyna Szatraj[1], Urszula Zielenkiewicz[1], Sebastian Piłsyk[1], Ewa Malc[3], Piotr Mieczkowski[3] , Joanna S. Kruszewska[1], Przemysław Bernat[4], Julia Pawłowska[2]

[1] *Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawinskiego 5A, 02-106 Warsaw, Poland*

[2] *Institute of Evolutionary Biology, Faculty of Biology, Biological and Chemical Research Centre, University of Warsaw, Zwirki i Wigury 101, 02-089 Warsaw, Poland*

[3] *High Throughput Sequencing Facility of UNC, Chapel Hill, NC, USA.*

[4] *Department of Industrial Microbiology and Biotechnology, Faculty of Biology and Environmental Protection, University of Lodz, Banacha Street 12/16, 90-237, Lodz, Poland*

Abstract

We describe the genomes of six Mucoromycotina fungi representing distant saprotrophic lineages within the subphylum (i.e. Umbelopsidales and Mucorales). We selected two *Umbelopsis* isolates from soil (i.e. *U. isabellina, U. vinacea*), two soil-derived *Mucor* isolates (i.e. *M. circinatus, M. plumbeus*), and two Mucorales representatives with extended proteolytic activity (i.e. *Thamnidium elegans* and *Mucor saturninus)*. We complement genome analyses with a description of their digestive capabilities, their cell wall carbohydrate composition, and total lipid profiles. Finally, we link the presence of endohyphal bacteria with observed characteristics.

One of the genomes, *Thamnidium elegans*, harbours a complete genome of an associated bacterium classified to *Paenibacillus* sp. This fungus displays multiple altered traits compared to remaining isolates regardless of their evolutionary distance. *T. elegans* has expanded carbon assimilation capabilities particularly efficiently degrades carboxylic acids, has a higher diacylglycerol: triacylglycerol ratio and phospholipid composition suggesting a more rigid cellular membrane.

Comparison of early-diverging Umbelopsidales with evolutionary younger Mucorales points at several differences particularly in their carbon source preferences and encoded carbohydrate repertoire. All tested Mucoromycotina shares features including the ability to produce 18:3 gamma-linoleic acid and fucose as a cell wall component.

Author Summary

In our paper, we report on the genomic sequences of six Mucoromycotina strains and an associated bacterium from *Paenibacillus* genus. Mucoromycotina are often studied in pathogenic context albeit their basic biology remains understudied. This manuscript expands on the collection of currently sequenced Mucorales and Umbelopsidales, including the first sequenced *Thamnidium* isolate, which was sequenced together with a *Paenibacillus* bacterium. The interaction with a bacterial partner alters the metabolism, cell membrane composition but not the exoskeleton of the fungus. The associated bacterium provided multiple enzymes that significantly expanded the digestive capabilities of the fungal host. Parallel sequencing and phenotyping of Mucorales and Umbelopsidales enabled us to look at the differences of both lineages within Mucoromycotina. We demonstrate that the predicted digestive capabilities are in line with experimental validation. Based on the cell wall composition data and genomic underpinnings of carbohydrate metabolism we were able to confirm the universal presence of fucose in Mucoromycotina cell walls. Fatty acid, phospholipid and acylglycerol composition support the usage of 18:3 gamma-linoleic acid as a chemotaxonomic marker of Mucoromycotina and corroborate TAG as a dominant storage lipid in these organisms.

Genomic features, digestive capabilities, fatty acid composition differ between Mucorales and Ubelopsidales pointing at subtle but significant changes in the course of Mucoromycotina radiation.

Introduction

Mucoromycotina subphylum comprises three orders: Umbelopsidales, Endogonales and Mucorales [1]. While Umbelopsidales and Mucorales group mostly saprotrophic fungi living in the soil, on dung and litter, Endogonales are known to establish symbiotic interactions with plants [2]. The ancestors of extant Mucoromycotina were among the first colonizers of land. These early-

branching fungi possess all of the traits commonly acknowledged as distinctive characteristics of the fungal kingdom such as apical growth, presence of ergosterol in the membranes and a cell wall made of chitin and beta-glucan [3]. Yet, their cell wall differs from ascomycetes and basidiomycetes by the presence of fucose and high amounts of N-acetylglucosamine and glucuronic acid [4,5] which is typical for other Opisthokonta rather than fungi. Mucoromycotina representatives form a fast-growing mycelium of haploid hyphae. The sexual phase includes the fusion of two gametangia and formation of a resting spore called zygospore [1].

Apart from decomposing organic matter as saprotrophs, some Mucoromycotina are capable of forming mutualistic and mycorrhiza-like associations with Haplomitriopsida liverworts [6,7]. Others are parasites of plants and animals [8,9]. The ecology of Mucoromycotina is poorly studied, which hinders understanding of their role in the ecosystem [10].

The best-studied order of Mucoromycotina is Mucorales. Its representatives are perceived mostly as a source of food spoilage, rotting crops and decomposing litter [11]. Due to their efficient plant matter degrading potential, Mucorales are used in various food production processes like the fermentation of soya beans known in the traditional Asian and African cuisine as tempeh [12] or cheese making [13]. Others are employed as bio producers of polyunsaturated fatty acids (eg. *M. circinelloides)*, β-carotene (eg. *Blakeslea trispora*, *M. circinelloides*, and *Phycomyces blakesleeanus*) and diverse hydrolases [14–16]. Certain Mucorales (e.g. *M. indicus*) are also able to produce ethanol [17] while others (e.g. *Rhizomucor miehei*) are commercially used for lipase production [18][11]. Mucorales lipid composition includes gamma-linolenic fatty acid 18:3 [19] which is found mostly in plants [20], algae [21] and not in Dikarya which produce the n-3 isomer of the C18 trienoic fatty acid alpha-linolenic acid.

Some Mucorales are involved in life-threatening, opportunistic infections with mortality rates reaching up to 40-80% [22]. There are general traits which predispose microorganisms to become

opportunistic pathogens, e.g. thermotolerance and ability to evade immune cells. Many non-pathogenic fungi are adapted to higher temperatures due to living in decomposing organic matter, often warmed up due to rotting processes [23,24], which allows them to survive inside the animal warm-blooded body.

Recently, endohyphal bacteria have been described in a few different fungal lineages of Mucoromycota (i.e. Mucoromycotina and 2 other closely related subphyla: Glomeromycotina and Mortierellomycotina). *Mycoplasma*-related endobacteria (MRE) were detected in Endogonales [25], Glomeromycotina [26], and Mortierellomycotina [27]. *Burkholderia*-related endobacteria (BRE) were identified in Mortierellomycotina, Mucorales and Gigasporales representatives [28]. *Paraburkholderia*-related endobacteria (PRE) were discovered in Mucorales, Umbelopsidales, and Mortierellomycotina representatives. Lately, information about two other endosymbionts of *Rhizopus microsporus* has been published: *Ralstonia pickettii* as an enabler of fungal invasion during infection [29] and *Stenotrophomonas* sp. whose role has not yet been established [30]. The activity of endohyphal and fungi-associated bacteria involve (but is not limited to) change in lipid metabolism [31–33], change in morphology of the fungus [31,32,34], control of the hosts' sexual [35], and asexual reproduction [36].

Here we describe the genomes of six Mucoromycotina species representing separated saprotrophic lineages within the subphylum (i.e. Umbelopsidales and Mucorales). Mucoromycotina are often studied in pathogenic context albeit their basic biology remains understudied. We selected two *Umbelopsis* isolates from soil: *U. isabellina* and *U. vinacea*, two soil-derived *Mucor* isolates: *M. circinatus* and *M. plumbeus*, and two Mucorales representatives: *Thamnidium elegans* and *M. saturninus* with proteolytic capabilities enabling them to colonize dung and animal substrate [37]. We complement genome analysis with a phenotypic description of their digestive potential, their cell wall carbohydrate composition, and total lipid profiles. Finally, we link the presence of endohyphal bacteria with observed characteristics.

Results

Genome assembly and annotation

Obtained assemblies showed diverse levels of fragmentation depending on genome size and abundance of repeats. Both *Umbelopsis* genomes assembled into fewer than 200 scaffolds whereas *Mucor* assemblies were significantly more fragmented (**Table 1**).

Genome completeness was verified using single copy fungal orthologous genes searched by BUSCO [38]. Only a few BUSCOSs were duplicated from 27 in *U. isabellina* up to 57 in *M. saturninus* suggests that sequenced genomes were not duplicated. The number of predicted genes, repeats and transposable elements (TEs) content correlates positively with genome size.

Table 1. Summary of genome sequence characteristics.

	Umbelopsis isabellina	*Umbelopsis vinacea*	*Mucor circinatus*	*Mucor plumbeus*	*Mucor saturninus*	*Thamnidium elegans*	*Paenibacillus* sp.
genome size [Mb]	22	23.2	55.7	49.5	40.3	31	6.9
genes	9030	9241	14148	11690	13020	10765	5954
tRNAs	129	134	194	274	262	191	119
TE number	37	59	1681	869	1166	541	-
% GC	42	43.09	31.89	31.1	36.59	33.15	44.61
% repeats	7.85	7.81	28.24	43.3	32.52	18	-
contigs	82	175	16245	8313	7254	3243	285
coverage (x)	60	68	26	28	31	61	8
N50	1335078	1328092	125937	49105	159134	100281	129168
L50	7	7	116	301	68	89	16
BUSCOS %	98.3	97.6	96.9	97.2	98.7	94.5	85.4

Repetitive elements

The largest group of TEs found in the genomes of Mucorales, with more than 200 copies per genome are class II elements belonging to Tc1/Mariner superfamily (**Supplementary Table S1**) but other DNA-repeats were also ubiquitous. EnSpm, PIF/Harbinger, hAT-Ac MuLE, Merlin, PiggyBac were present from one up to a hundred copies in all sequenced Mucorales. EnSpm,

PIF/Harbinger, hAT-Ac and MuLE are ubiquitous in fungi in general, whereas Merlin elements are characteristic for basal fungal lineages and were apparently lost in Dikarya [39]. Unlike Mucorales, Umbelopsidales genomes had just a few copies of Tc1/Mariner, hAT-Ac, and a single Ginger element. Class I retrotransposon landscape is dominated by LTR retrotransposons from Ty3/Gypsy and LINE/L1 elements which were found in all genomes except for *U. vinacea* which lacks the L1s. Surprisingly the omnipresent Ty1/Copia elements were identified only in the genome of *M. plumbeus* (two copies). Also, all sequenced species had LTR/DIRS elements, characteristic for their tyrosinase integrase and are absent from Dikarya [40]. Mucor genomes harbour also numerous Helitrons, known for their rolling circle replication mechanism. The presence of Helitrons has been observed in other *Mucor* species previously [41]. Noteworthy, Helitrons often hijack neighbouring genes and are efficient vectors of HGT in other fungi [42]. Overall, *Umbelopsis* genomes contained ten-fold less transposable elements per genome compared to Mucorales, which is expected taking into account the differences in genome size.

RNAi and other defence mechanisms

Fungal genomes are usually protected *via* diverse mechanisms that include fungal repeat-induced point mutation (RIP), methylation induced premeiotically (MIP), meiotic silencing of unpaired DNA (MSUD) and quelling. It has been described that Mucormycotina possesses RNAi pathway components [43] which are also important for fungus biology [44,45]. We scanned the sequenced genomes with HMM profiles of RNAi core enzymes, namely Dicer, Argonaute and RdRP, and demonstrate that all are present in all six fungal genomes, with duplication of RdRP in all Mucorales (from three to five copies per genome) and a single copy in both *Umbelopsis* isolates (**Supplementary Table S2)**. *U. vinacea, M. circinatus, M. plumbeus* harboured duplicated Argonaute proteins.

Self-non-self and the fungal immune system based on Nucleotide Oligomerization Domain (NOD)-like receptors (NLRs) could be identified neither in the six genomes described in this work nor in 14 predicted Mucormycotina proteomes available at NCBI. The typical NLR central domains NACHT or NB-ARC seem to be absent from those genomes.

Detection of *Paenibacillus* bacteria in *Thamnidium* genome

Initial *Thamnidium* assembly contained genome fragments characterized by two clearly distinct GC content ratios and for that reason was additionally analysed as a metagenome. It is composed of two easily distinguishable fractions, one belonging to a presumed fungal host and the latter to an associated bacterium representing Paennibacillales, Firmicutes. Fungal and bacterial genomes were re-assembled separately. Remaining five Mucoromycotina assemblies did not contain significant amounts of sequences with sequence similarity to non-Mucoromycotina taxa.

Paenibacillus features

The bacterial genome is moderately complete (BUSCO score 85.7%) and encodes 5815 genes. Its closest relative, *Paenibacillus* sp. 7523-1, despite DDH similarity above 70% threshold for two out of three DDH calculation formulas implemented in TYGS (67.2; 88.3; 72.9), differs in GC content by 2.35% which supports the separation of the newly sequenced strain as a new species (**Supplementary Figure S1, Supplementary Table S11**). A phylogenetic tree inferred from 16S RNA of related isolates shows its proximity to *P. illinoensis* isolates despite differences in GC content (**Fig.1.**). GC content of the identified *Paenibacillus* genome could be altered by misplaced fungal reads. Despite discarding all reads mapping on eukaryote genomes when assembling the *Paenibacillus* genome it is still possible that some fungal fragments, not similar to other eukaryotic sequences, could not be filtered out.

The bacterial genome has only one complete antimicrobial locus, a rifamycin-inactivating phosphotransferase RphD, identified by scanning using ABRicate. We found a partial beta-lactam

resistance pathway lacking the central beta-lactamase (four copies of beta-N-acetylhexosaminidase, BlaI family transcriptional regulator, oligopeptide transport system ATP-binding protein from oppA to oppF, penicillin-binding protein 1A and 2A). Additionally, the *Peanibacillus* genome encodes a single M56 peptidase and 15 proteins bearing an S12 domain. Family M56 includes BlaR1 which is the antirepressor in beta-lactamase operon [46]. Family S12 groups carboxypeptidases B vital for cell wall synthesis and remodelling, as well as beta-lactamases [47]. Taken together, all these traces point at a potential beta-lactam resistance of *Paniebacillus*. Glycopeptide antibiotic resistance may also be present in some form since partial vancomycin resistance operon was also identified, once more only regulation and accessory components are preserved. The possible activity of these proteins and their relevance for resistance is not known.

The *Paenibacillus* genome harbours vitamin biosynthesis pathways e.g. cobalamin (B12), riboflavin (B2), menaquinone (K2) and thiamine (B1). It also has all the necessary genes for molybdenum cofactor synthesis and manipulation. *Paenibacillus* has many quorum sensing genes and those coding for flagellar and mobility proteins. Its metabolic potential is described below in parallel with its fungal host and remaining fungal genomes.
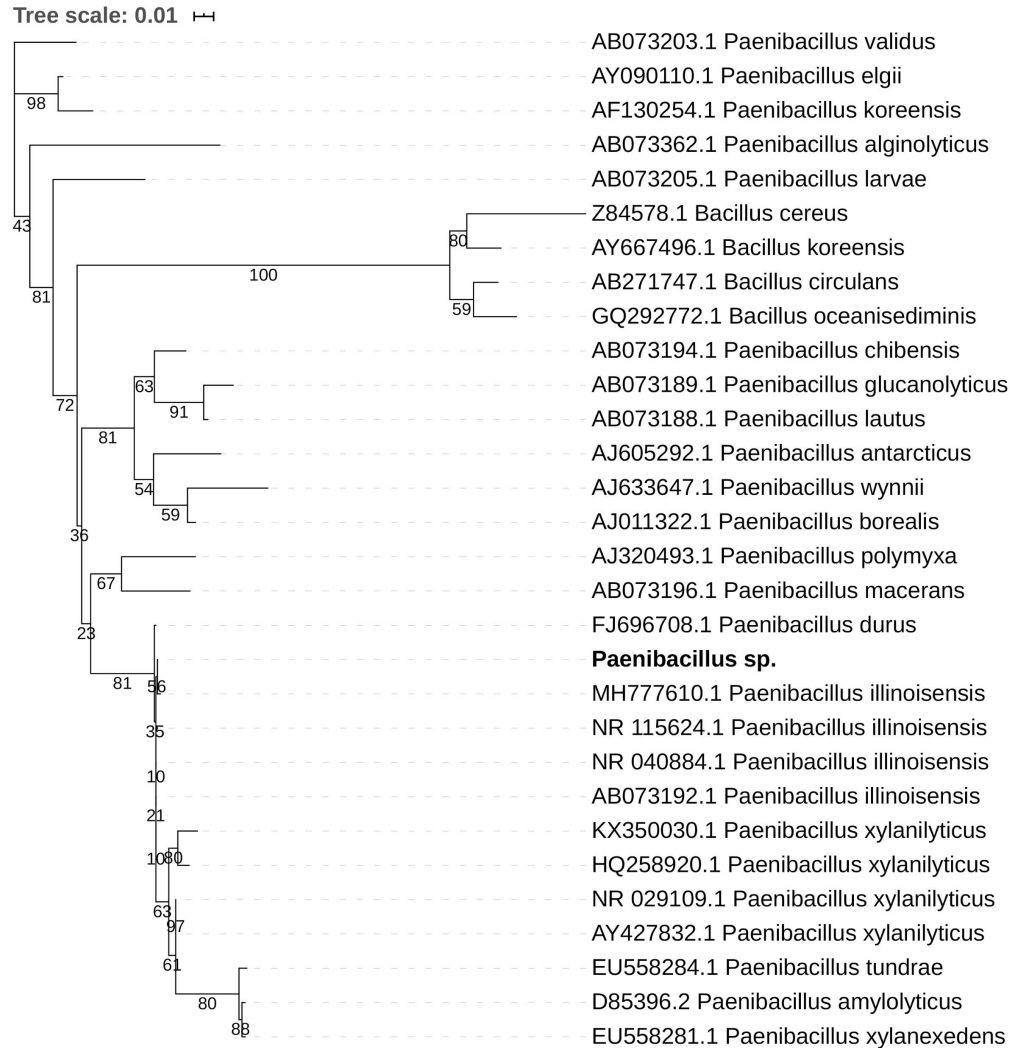
Figure 1. Maximum likelihood phylogenetic tree of 16S rRNA shows the sequenced strain (bold font) groups with *P. illinoisensis* isolates. Branches with support values below 50 were collapsed, the tree was rooted with *Bacillus* sequences.

The *Paenibacillus* genome encodes 3 copies of Peptidase_U32 (PF01136), a collagenase that may facilitate meat degradation by *Thamnidium elegans* which is one of the few fungi known for this property [48]. This protein is found mostly in Firmicutes and is absent from eukaryotic genomes. The U32 collagenases are considered as virulence factors in animal-infecting bacteria [49]. The U32 collagenase (PrtC) together with urease subunit alpha (UreB) are parts of the *Helicobacter pylori* arsenal used in epithelial cell invasion [50]. *Paenibacillus* sp. genome harbours

all three urease subunits (alpha, beta and gamma). Moreover, we identified genes coding other urea processing enzymes including: cyanuric acid amidohydrolase, biuret amidohydrolase (BiuH), urea carboxylase and two copies of allophanate hydrolase (AtzF). The *Paenibacillus* genome encodes also one copy of ulilysin Peptidase_M43 (PF05572) with possible gelatinase function [51], and 33 copies of Peptidase_M23 (PF01551), which includes mostly bacterial peptidoglycan hydrolases but also prokaryotic collagenases [52].

Genes coding for CAZymes, proteases and transport-related proteins

Proteases

The overall profile of encoded proteases is very similar in all sequenced fungi (**Fig.2.**), with high numbers of encoded pepsin-like A01A peptidases, FtsH-like M41 peptidases, M48 peptidases active on di- and tripeptides, proteasome peptidases T1, lysosomal peptidases C26, and ubiquitin-specific proteases C19. These proteases, except for pepsin, contribute to intracellular protein turnover and regulation. Pepsin and subtilisin proteases are found in high copy numbers in all of the genomes pointing at a high degradation potential of these saprotrophic organisms. In all analyzed fungal genomes, we found an expansion of the C44 family. Proteins from this family are homologs of glutamine-fructose-6-phosphate transaminase (GFPT) and its precursor. GFTP is known to control the flux of glucose into the hexosamine pathway which plays a crucial role in the regulation of chitin synthesis in fungi [53]. There is a two-fold expansion of S8A serine proteases in Mucorales when compared to Umbelopsidales (**Supplementary Table S3**). The same observation applies to family I4 which groups inhibitors of S8 peptidases. Umbelopsidales possess several families of cysteine peptidases (C15, C40, C110) and metalloproteases (M14A and M20A) absent from Mucorales genomes. The metalloproteases are likely involved in protein degradation whereas the cysteine peptidases have no obvious function in fungi.

Among sequenced Mucorales, the average number of protein-coding genes per peptidase family is elevated, which could be a consequence of the whole genome duplication (WGD) described in *Mucor* and *Phycomyces* by Corrochano and coworkers [54]. However, BUSCO results show a low level of duplicated genes in all of the isolates. The genome of *Thamnidium* stands out as it encodes fewer proteases than remaining Mucorales. However, it may benefit from enzymes provided by its endohyphal bacterium which has genes encoding meat crumbling collagenase U32 and plenty of typically bacterial proteases representing families A36, A25, S15, S55, S66, S51, M29, and U57.
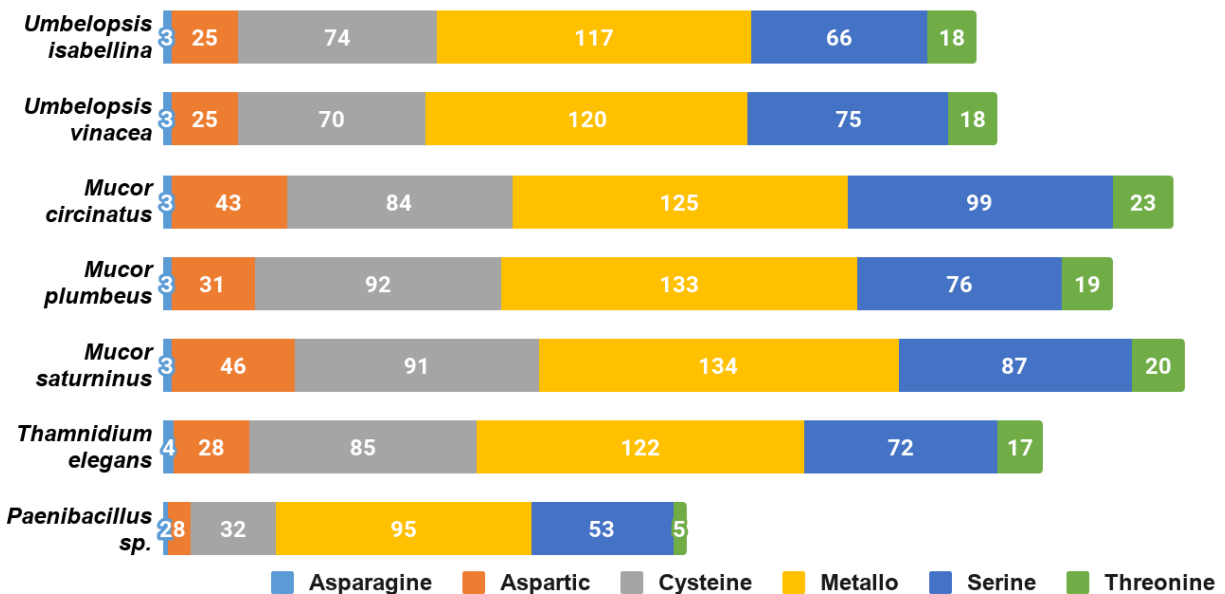

Figure 2. Distribution of genes encoding particular catalytic types of proteases among sequenced isolates.

Carbohydrate active enzymes

All six genomes encode an extensive repertoire of carbohydrate-active enzymes (**Fig.3.**). A high number of glycohydrolases is present exclusively in the genomes of *Umbelopsis* spp. and absent from all four Mucorales; these are GH1, GH2, lysozymes (GH24 and GH25), bacterial β-

11

xylosidase GH52, endo-β-1,4-galactanase GH53, β-1,3-glucanase GH55, GH71, rhamnosidase GH78, α-mannosidase GH92, exo-α-L-1,5-arabinanase GH93, α-L-fucosidase GH95 and peptidoglycan lytic transglycosylase (GH104), enzymes that hydrolyze fructose-containing polysaccharides such as inulinases (GH32), endoglucanase like cellulase and endomannanase (GH5), phosphorylases (GH65), β-glucuronidases (GH79) and glucuronyl hydrolases (GH88) (**Supplementary Table S4)**. All studied fungi have numerous copies of enzymes potentially involved in nutrient degradation and metabolite modification, for instance: α-amylase (GH13), β-amylase (GH14), feruloyl esterase/acetyl xylan esterase (CE1), cholinesterase (CE10). All studied genomes have also an acetyl esterase that liberates acetic acid from xylo-oligomers (CE16). All also have homologs of GH13 α-amylases, but only Umbelopsidales have representatives of the GH13_40 subfamily. This subfamily seems to group exclusively eukaryotic α-amylases. Surprisingly, *Umbelopsis* spp. have also PL12 lyase, a bacterial enzyme used to access metazoan debris by specifically cleaving heparan sulfate-rich regions of acidic polysaccharides which are found in connective tissues [55]. Other typically bacterial enzymes found here are β-xylosidase GH52 and peptidoglycan lytic transglycosylase GH104. This repeated presence of bacterial enzymes may be a result of a horizontal gene transfer event from either endohyphal or free-living bacteria to Umbelopsidales or its ancestor. We also observed an expansion of GH16 hydrolases and hexosaminidase (GH20) in Umbelopsidales compared to Mucorales. On the other hand, *Umbelopsis* spp. lack multiple enzymes belonging to several Cazy families, like acetyl xylan esterase (CE6), cellulose-binding modules (CBM6 and CMB9), chitin-binding module (CBM14), cell wall-related α-1,6-mannanase (GH76) and endo-β-1,4-mannanase (GH134).

Noteworthy, all six fungi have expansions of protein families involved in cell wall remodelling. These include: chitooligosaccharide deacetylase (CE4), chitinase (GH18), chitin synthase (GT2), α-1,2-mannosyltransferase (GT15), β-1,3-N-acetylglucosaminyltransferase / β-1,3-

glucuronyltransferase (GT49). Additionally, we found expansions of chitin or peptidoglycan (CBM50) binding modules and additional, broad range binding module (CBM13). Protein glycosylases are also particularly abundant in sequenced taxa and include homologs of Drosophila O-fut2 belonging to protein O-α-fucosyltransferase (GT68) and glycoprotein α-1,3-L-fucosyltransferase (GT10). The latter is encoded by 3 *Mucor* taxa and other known Mucoromycotina but not by *Umbelopsis* spp. Fucosylation may additionally mask Mucorales against animal immune systems, as fucose is widely present in the blood.

NagA α-N-acetylgalactosaminidase, NagZ β-N-acetylhexosaminidase (GH109) involved in peptidoglycan recycling occur in multiple copies not only in *Paenibacillus* genome but also in sequenced fungal genomes. It is worth noting that *Peanibacillus* has three times more copies (30) of Nag genes than studied fungi. No proteins from this family had been characterized in Eukaryota.

Thamnidium elegans possesses the narrowest set of carbohydrate-processing enzymes which, like in the case of proteases, might be expanded by its associated bacterium. *Paenibacullus* encodes several unique enzymes, like pectate lyases (PL1, PL10, PL9) and rhamnogalacturonan (PL11), oligogalacturonate (PL22) and L-rhamnose-α-1,4-D-glucuronate lyases (PL27). It also encodes lytic chitin monooxygenase (AA10) which may be either used to manipulate the host or to acquire nutrients. Additionally, *Paenibacillus* has expanded digestive repertoire including xylanases and xylosidases (GH11, GH30, GH43) and galactosidases (GH4, GH42, GH110). Interestingly, it encodes bacterial lipid-A-disaccharide synthase (GT19) and murein polymerase (GT51). *Paenibacillus* genome harbours 39 copies of S-layer homology (SLH) domain proteins which are important for biofilm formation [56].
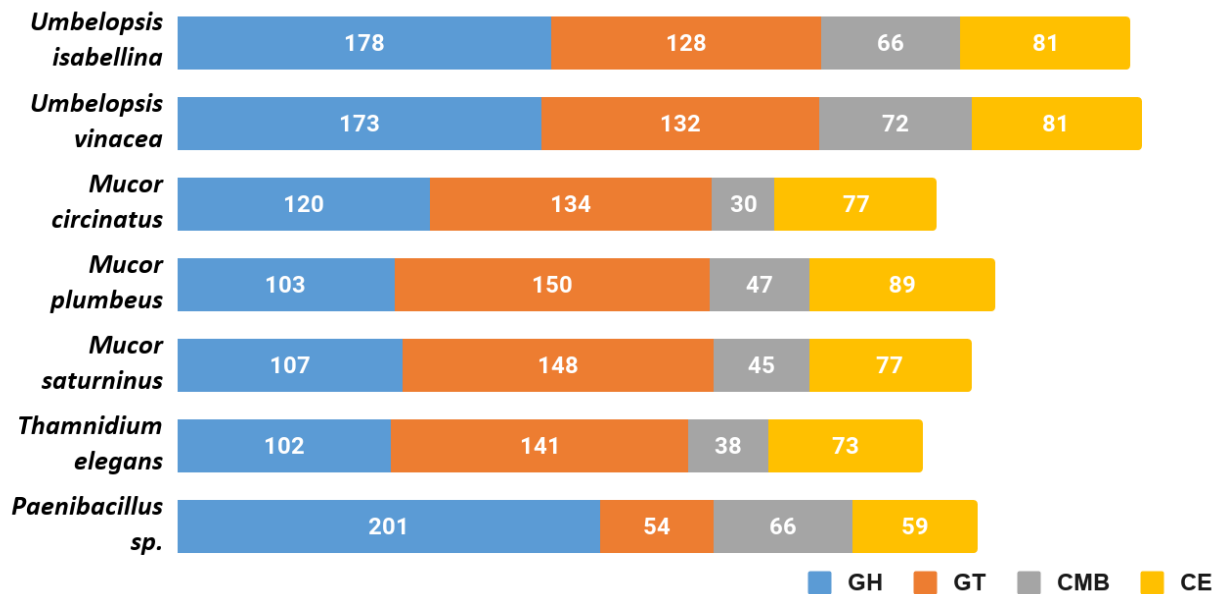
Figure 3. Cazyme distribution in the sequenced genomes. Abbreviations: GH-glycoside hydrolases, GT-glycosyltransferases, CMB-carbohydrate-binding modules, CE-carbohydrate esterase.

Transporters

Sequenced Mucorales have several classes of transporters: Amino Acid-Polyamine-Organocation (APC), Glycoside-Pentoside-Hexuronide (GPH), Putative Tripartite Zn2+ Transporter (TZT), ATP-binding Cassette (ABC), Major Facilitator Superfamily (MFS) and Drug/Metabolite Transporter (DMT) (**Supplementary Table S5)**. *Mucor circinatus* possesses a huge expansion of MFS (347 compared to a median of 209 copies per genome). *Thamnidium elegans* has fewest transporter-encoding genes, but it is again possible that this is an effect of relationship with bacteria serving as a reservoir of necessary enzymes. Both *Umbelopsis* representatives have fewer transporters than analysed Mucorales.

Metabolic clusters, secondary metabolites and cofactors

The genome of *Paenibacillus* harbours NRPS, terpene, bacteriocin, T3PKS and S-layer-glycan producing clusters. Mucoromycotina were long considered devoid of secondary metabolite clusters. A review by Kerstin Voigt and colleagues [16] showed genetic determinants for natural product synthesis present in all analysed genomes. Our newly sequenced genomes encode between 3 and 8 secondary metabolite clusters (according to AntiSMASH scans) belonging to different classes. Interestingly, all of the genomes encode terpene clusters (**Supplementary Table S6)** which potentially could produce new natural products like those isolated from *Mortierella* [57].

Umbelopsidales additionally have a trans-2,3-dihydro-3-hydroxyanthranilate isomerase (PhzF) involved in phenazine biosynthesis with yet unknown biological product in fungi [58]. This gene is also present in other fungi and in Endogonales but seems to be missing in Mucorales. Umbelopsidales produce also a citronellol/citronellal dehydrogenase which converts citronellol to citronellic acid an odorous compound with antimicrobial properties. All four Mucorales isolates have a single gene coding for salicylate hydroxylase which is involved in plant host manipulation by *Epichloë* [59] and other fungi.

All analysed genomes have a B12 dependent methionine synthase homologous to the animal variant of the enzyme (PF02965, Met_synt_B12) and B12 independent synthase (PF01717, Meth_synt_2) encountered in Dikarya. Both enzymes have recently been found in other Mucoromycotina [60] and their presence suggests at least partial B12 dependence. Moreover, *Umbelopsis* spp. have methylmalonyl-CoA mutase (MUT), methylmalonyl-CoA/ethylmalonyl-CoA epimerase (MCEE), cob(I)alamin adenosyltransferase (MMAB, pduO) and ribonucleoside-triphosphate reductase (rtpR) - all using cobalamin. Mucorales have a B12-independent ribonucleotide reductase class II (nrdJ) and lack methylmalonyl-CoA mutase and MUT-epimerase genes.Sex locus

Sequenced isolates belong to heterothallic genera [64] and genomic screening shows a single sex locus per genome. Generally, these loci contain a single high mobility group (HMG)-domain transcription factor gene (*sexP* or *sexM*), flanked by genes for an RNA helicase (*rnhA),* a triosephosphate transporter (*tptA*) and an alginate lyase (*agl*) [65].

The sex locus of *T. elegans* and *M. plumbeus* is organized like in other Mucorales with genes in the following order *agl/tptA/sex/rnhA* [65]. *M. saturninus* lacks the *tptA* gene and has a sex locus architecture (*agl/sexP/rnhA)* like *M. mucedo* [66] . *M. circinatus* has remains of an integrase instead of the *sex* gene between *tptA* and *rnhA* genes *(agl/tptA/rve/rnhA)*. Mobile element insertions in the sex locus have already been documented in *Phycomyces blakesleeanus* [67].

The *rnhA* gene seems to be excluded from the sex locus in both Umbelopsidales genomes, in which the sex locus is organized like in Mucorales but followed by a gene with an additional protein of unknown function belonging to DUF2405 (PF09435) family (*agl/tptA/sexP/DUF2405). S*chultz and colleagues described a similar architecture with the DUF2405 for *Umbelopsis ramaniana* from JGI database [68].

Carbon assimilation profiles

Carbon assimilation profiles obtained for six Mucoromycotina strains by screening on Biolog FF microplates are summarized in Supplementary Table (**Supplementary Table S7)**. None of the analysed strains was able to use a full set of 95 tested carbon sources. Each strain was able to grow on 40 to 70 different substrates and had a unique carbon assimilation profile (**Fig.4**) with *T. elegans* being the most versatile degrader (**Fig.5**).
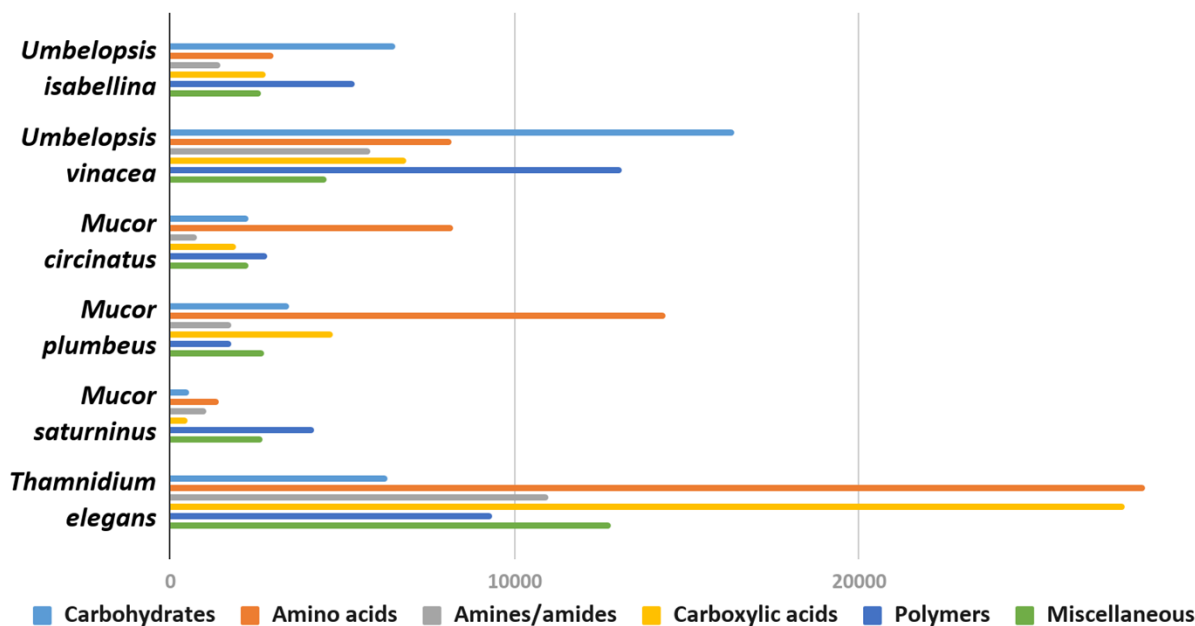
16

Figure 4. Usage of different types of carbon sources by particular isolates, mean area under the curve (AUC) values aggregated for guilds of carbon sources (according to Preston- et al. [69]).

Umbelopsis spp. were more efficient in the utilization of carbohydrates, while Mucorales representatives showed the fastest growth rate on amino acids. *Umbelopsis* fungi were able to utilize Adonitol, d-Galacturonic Acid, Maltitol, β-Methyl-D-Glucoside and d-Raffinose whereas none of the four Mucorales grew on these substrates. Additionally, similarly to *Mortierella elongata* AG77 (Uehling et al. 2017), they represented elevated growth rate on other carbohydrates, like d-Galactose, d-Mannose, l-Arabinose, l-Rhamnose, d-Trehalose and lipid (Tween 80), which is consistent with predicted gene models (see the chapter on CaZymes - the presence of rhamnosidase GH78, α-mannosidase GH92, exo-α-L-1,5-arabinanase GH93, α-L-fucosidase GH95). All four *Mucor* species can utilize D-Malic Acid and L-Malic Acid while these substrates seem inaccessible for *Umbelopsis* spp., which have only 3 copies of lactate dehydrogenase whereas *Mucor* spp. have from four to six enzymes from this family. Only *T. elegans* was able to utilize m-Inositol, Sedoheptulosan, β-Hydroxy-butyric Acid, p-Hydroxyphenyl-acetic Acid and L-Threonine. Additionally, *Thamnidium* is distinguished by very efficient development on carboxylic

acids. This ability may be explained by the presence of endobacteria whose proteome harbours representatives of M14 and M20 carboxypeptidase families. M14A is present uniquely in both *Umbelopsis* genomes while M14C occurs only in *Paenibacillus*. Also, peptidase T (M20B) is present exclusively in *Paenibacillus*, whereas fungal genomes have several copies of the remaining M20 subfamilies. Some of the analyzed compounds are available only to *Paenibacillus*, based on genomic evidence, e.g. 3-oxoacid CoA-transferase missing from sequenced Mucorales is present in Dikarya and bacteria, including *Paenibacillus*.
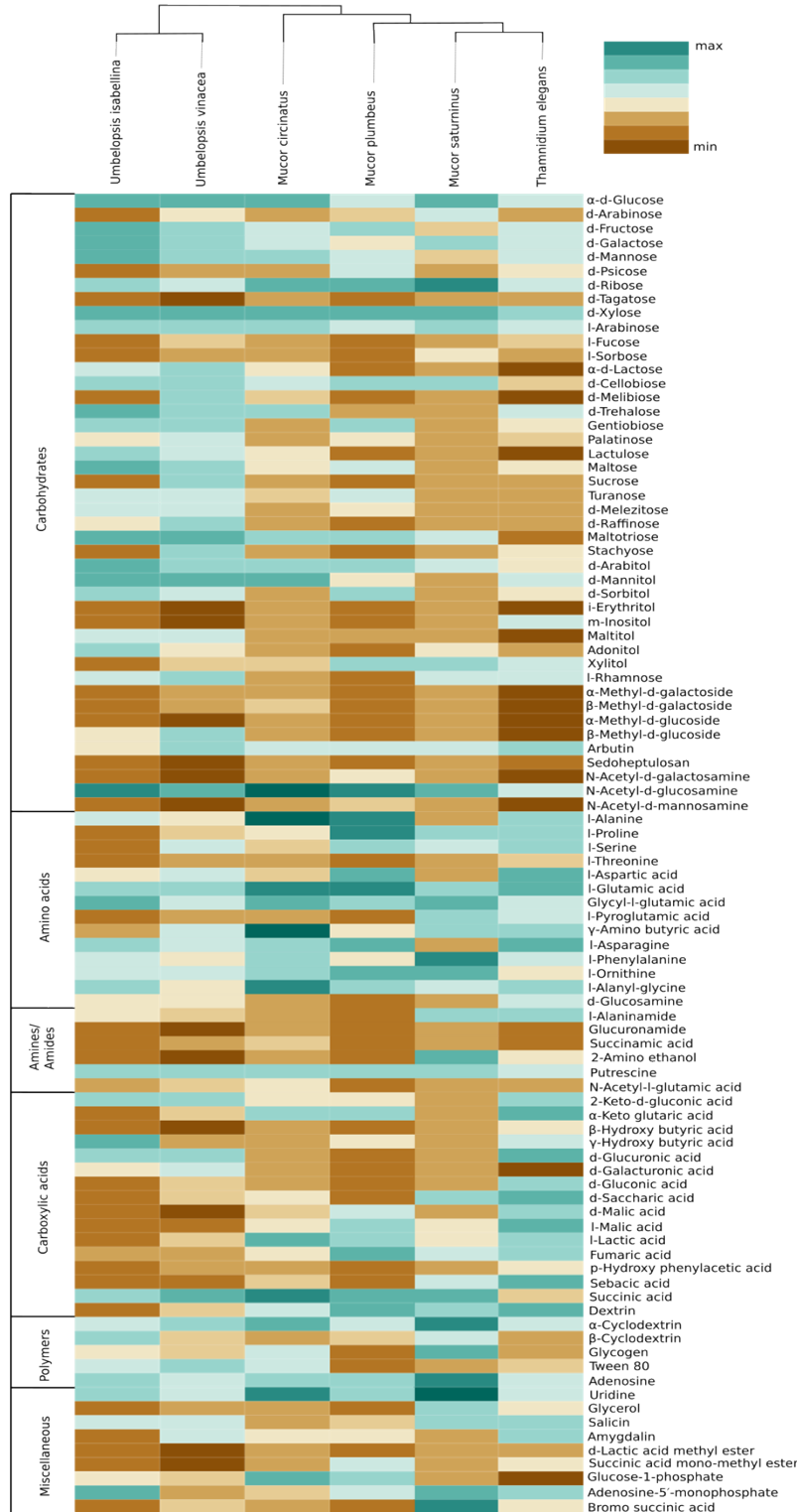
Figure 5. Heatmap representing carbon source utilisation capacity of 6 Mucoromycotina representatives obtained from Biolog FF MicroPlates (scale from brown - not used, to dark

green - used very efficiently). The phylogenetic tree is illustrating evolutionary relationships of tested fungal strains. Tested carbon sources are grouped into guilds according to Preston-Mafham et al. [69].
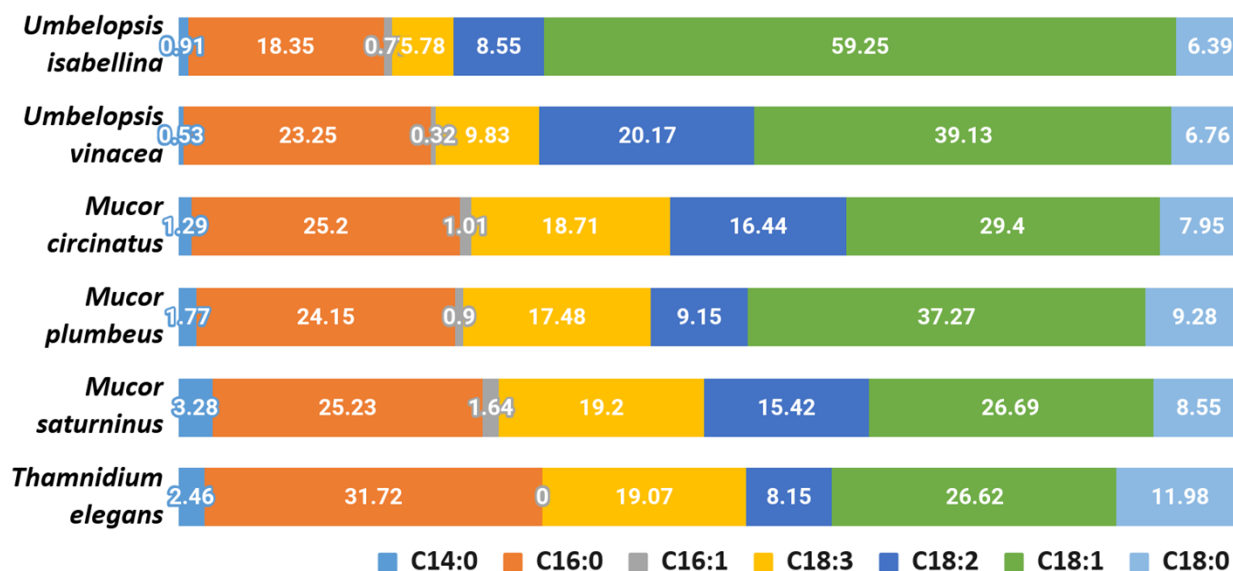
Fungal lipids

Figure 6. Composition of fatty acids in the analysed strains

For all six Mucorales, we determined the composition of sterols, fatty acids and phospholipids (**Supplementary Tables S8, S9 and S10)**. The presence of ergosterol, an important component of fungal plasma membranes, contributing to their stability, was confirmed in all 6 fungal biomasses.

Analysed strains contained the following fatty acids: saturated (14:0, 16:0, 18:0), monounsaturated (16:1, 18:1), and polyunsaturated (18:2, 18:3) (**Fig.6, Supplementary Table S8**). The lowest levels were observed for myristic (14:0) and palmitoleic (16:1) acids. A high percentage of oleic acid (18:1) is characteristic for the *Umbelopsis* strains and has also been reported by others [70–72]. It is also the dominant fatty acid for all isolates except for *Thamnidium*

20

elegans. A higher share of linolenic acid (18:3) was observed in Mucorales 18.5% compared to an average of 7.5% in Umbelopsidales. However, the efficiency/quantity of the production/synthesis of particular fatty acids depends, inter alia, on the type of substrate used [73]. Among examined strains, *T. elegans* profile of fatty acids showed the lowest unsaturation index with the domination of palmitic acid (16:0, 31.7%) and the highest share of stearic acid (18:0, 12%), which may suggest increased membrane stiffness and its lower fluidity. Paradoxically, *T. elegans* has more desaturases than the remaining five isolates (6 compared to 4). Production of unsaturated fatty acids involves specialized desaturases. The delta-6-fatty acid desaturase involved in the conversion of 18:2 to 18:3 fatty acid and delta-12-fatty acid desaturase processing 18:1 to 18:2 were experimentally characterized in diverse Mucoromycota including *T. elegans, Mortierella alpina* and *Rhizopus oryzae* [74–76]. All isolates had both delta-6- and delta-12-fatty acid desaturases. *T. elegans* has four copies of delta-12-fatty acid desaturases with an N-terminal short accessory domain of the unknown function (PF11960, DUF3474) which often occurs with desaturase domains and a single delta-6 desaturase with a Cytochrome b5-like Heme/Steroid binding domain (PF00173.28 Cyt-b5).

Using LC-MS/MS we determined the composition of phospholipids (PLs) present in the biomass of analysed fungi (**Fig.7**, **Supplementary Table S9**). We identified 84 PLs species belonging to 6 classes: phosphatidic acid (PA), phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), phosphatidylinositol (PI), and phosphatidylserine (PS). Sixty-two of them were chosen for subsequent quantitative assessments. This analysis of the PLs for selected strains revealed that PC and PE were predominant for selected Mucoromyotina and constituted up to 56% and 41% of the total cell PLs, respectively. PA, a lipid signal which usually constitutes a minute portion of PLs, in *T. elegans* and *M. saturninus* was found at 8% level. According to several reports, the increased levels of PA in living cells are a consequence of biotic and abiotic stress [77,78]. The open question remains whether *T. elegans* and *M. saturninus* experienced

stressing conditions during colony growth in the lab, or the observed increased level of PA is native for them as an adaptation to grow on animal substrate.

An important parameter describing the physical characteristics of biological membranes is PC/PE ratio. PC has a bigger head group than PE. Tighter packing of PE and its acyl chains negatively influences fluidity of membranes as opposed to PC [79]. PC can be synthesized from PE, so they are closely related, and the balance between PE and PC is crucial for maintaining the physiological structure of the cell membrane. Among selected species, *T. elegans* showed the lowest PC/PE ratio (0.79), which may indicate a relatively greater stiffening of the membrane compared to other strains, which is consistent with lowered fatty acids unsaturation index.

Moreover, it is worth mentioning, that only *T. elegans* and *U. vinacea* had traces of PE 16:0 16:1 which is characteristic for diverse bacteria, including isolates of *Pseudomonas putida, Bacillus subtilis* or *Escherichia coli* [80–82]. The dominant PLs species for all selected fungal strains were PC 18:2 18:3, PC 18:1 18:3, PC 18:1 18:1, PE 18:1 18:3, and PE 18:1 18:1. Similar results have already been reported for the PL profile of another Mucoromycotina representative, *Cunninghamella elegans* [78].

Other lipids which are important for the fungal metabolism are acylglycerols, triacylglycerols (TAGs) and diacylglycerols (DAGs) (**Fig.8**, **Supplementary Table S10**). Yeasts store lipids mainly in the form of TAGs, and to less extent, DAGs [33]. In all six fungal strains, both acylglycerols were found. All analysed strains accumulated more TAGs than DAGs but in *T. elegans* this ratio was significantly shifted towards 3:1 TAG to DAG, in contrast to the 5:1 in *M. plumbeus* or even 99:1 in *M. saturninus*. Since TAG synthesis in fungi requires PA and DAG, both present in *T. elegans*, observed DAG/TAG ratio might be a consequence of inhibition of acyl-CoA-dependent diacylglycerol acyl-transferase (DGA1) [83].

Lastovetsky and co-workers [33] reported that PE/TAG ratio plays a crucial role in establishing and maintaining symbiosis of the fungus *Rhizopus microsporus* (Mucoromycotina) and its *Mycetohabitans* endobacteria. PE/TAG ratio close to 1:1 is characteristic for symbiosis and departure from that balance might shift the interaction towards antagonism. Interestingly, it was observed that among six strains, the PE/TAG ratio for *T. elegans* was closer to 1:1 (2.21) compared to the average (6.4) for all tested fungi. Analysed strains all contain several copies of diacylglycerol kinase DGK genes (3-5) which is deemed responsible for maintaining the balance between TAG and PE.

The genome of *Paenibacillus* sp. encodes processive diacylglycerol beta-glucosyltransferase required for the synthesis of beta-diglucosyl-DAG - a predominant glycolipid found in Bacillales [84], as well as the Ugp sn-glycerol-3-phosphate transport system which transports glycerol-3-phosphate, essential for phospholipid biosynthesis.
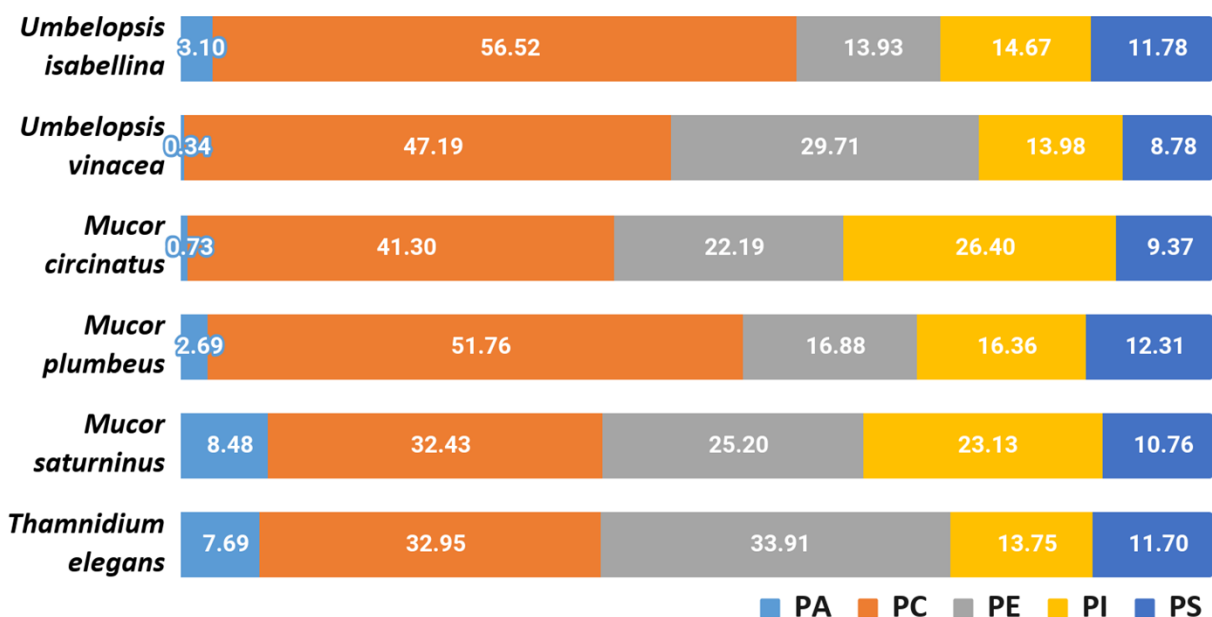
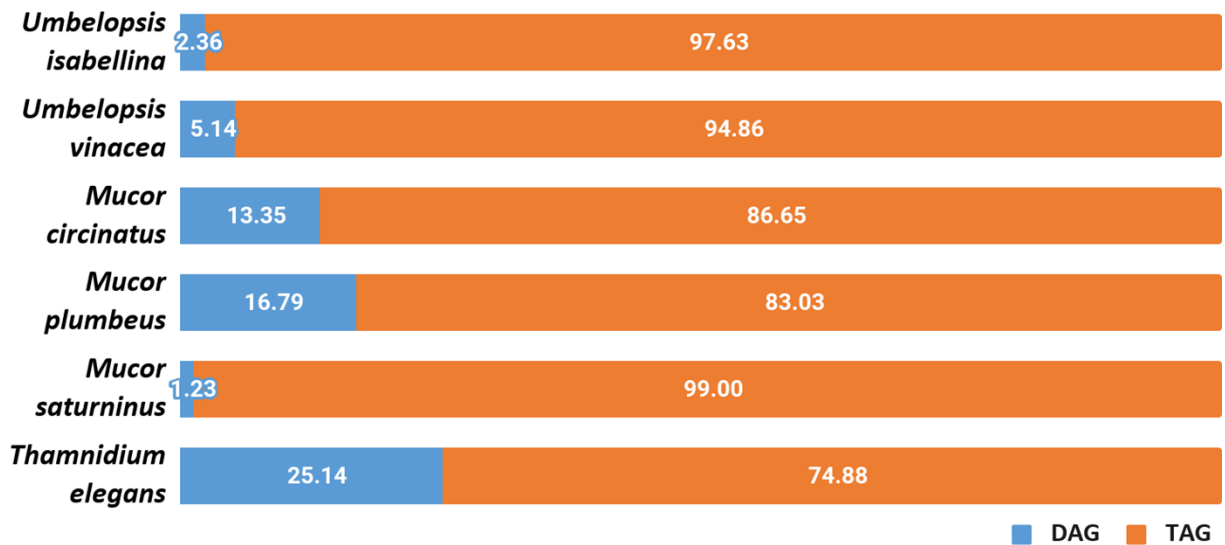**Figure 7.** Composition of phospholipids in the analysed strains

Figure 8. Distribution of diacylglycerols and triacylglycerols in the analysed strains

Cell-wall carbohydrates

A quantitative analysis of the cell wall carbohydrates revealed the presence of high amounts of glucosamine and fucose, and low amounts of mannose, galactose, and glucose compared to an ascomycetous fungus *Trichoderma reesei* TU-6 (**Fig.9**). Glucosamine content in their cell wall was up to 10 fold higher compared to *Trichoderma*. High glucosamine fraction can be a highlight of the chitin-chitosan cell wall typical for Mucoromycotina [5].
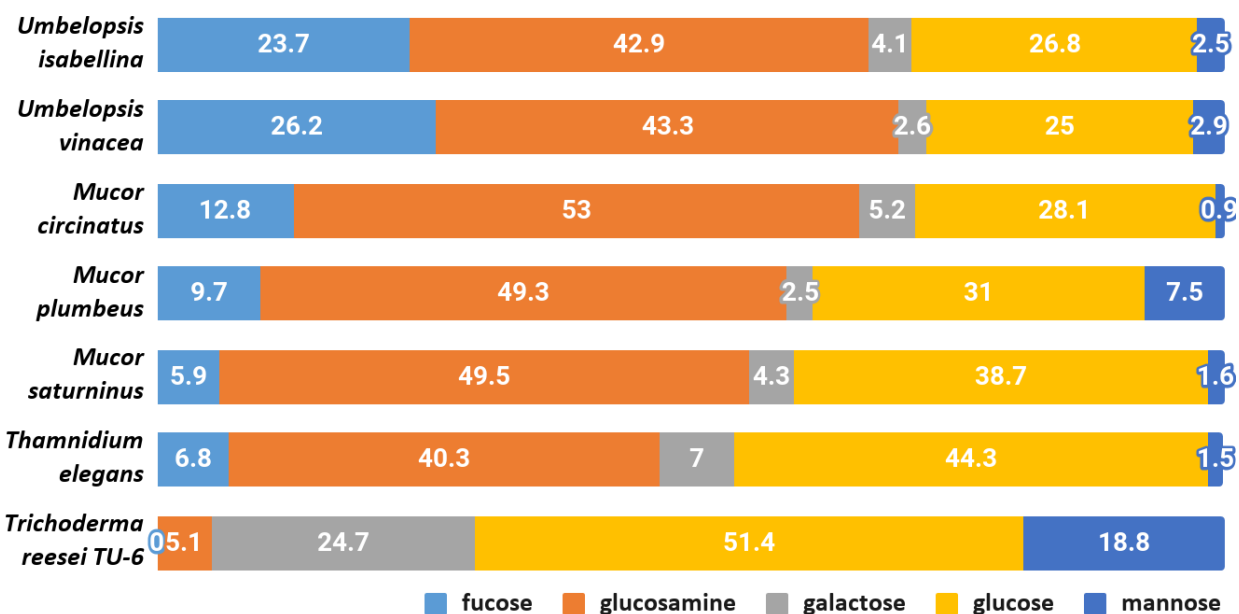

Figure 9. Composition of cell-wall carbohydrates.

Chitin together with glucans participates in the rigidity of the cell wall. The total content of glucosamine (representing chitin/chitosan fraction) and glucose (representing glucan fraction) in four Mucorales: *M. plumbeus*, *M. saturninus*, *M. circinatus* and *T. elegans* was above 80% of the cell wall carbohydrates. The amount of these two carbohydrates in *Umbelopsis* strains was lower and reached about 66 and 71% in *U. isabelina* and *U. vinacea*, respectively. *Trichoderma* had only 55% of these sugars in the cell wall and, in addition, glucose was the dominant one.

Melida and coworkers [4] reported that genes encoding chitin biosynthesis (CHS) and its modification (CDA) were present in more copies in *Phycomyces blakesleeanus* and *Rhizopus oryzae*, compared to *Neurospora crassa* (Ascomycota), as well as fewer genes encoding glucan synthases. These results were confirmed in other studies on Mucorales [85] and our results show that respective CaZymes are abundant both in sequenced Umbelopsidales and Mucorales representatives. This could explain the possibility of synthesizing a huge amount of chitin in Mucoromycotina strains compared to *Trichoderma*.

Furthermore, all Mucoromycotina contain fucose, a carbohydrate characteristic for this subphylum, especially in such a high amount. Our study revealed that two *Umbelopsis* strains had from twice to three-fold more fucose when compared to *M. circinatus* and *M. saturninus*, respectively. Presence of fucan in the cell wall was previously reported for *P. blakesleeanus* and *R. oryzae* [4]. During the synthesis of fucan, fucose is transferred from GDP-fucose to polysaccharides by α-fucosyltransferase and this enzyme was detected in the membrane fraction of *M. circinelloides* and partially characterized [85].

The α-fucosyltransferase encoding genes were found in 2 and 4 copies in the genomes of *P. blakesleeanus* and *R. oryzae* respectively, and they were not detected in *N. crassa*, which is in accordance with the fact that no fucose was detected in the cell wall of *Neurospora* [4] and *Trichoderma* (this study). Differences in the copy number of α-fucosyltransferase encoding genes do not explain the significant differences in the amount of fucose in the cell wall of *Mucor* spp. (4-6 copies) and *Umbelopsis* spp. (3-4 copies). All six genomes have 1-3 copies of GDP-L-fucose synthase.

Thamnidium elegans despite the associated bacteria shows a cell wall carbohydrate profile similar to remaining Mucorales and other Mucoromycotina. This suggests that the interaction with a

bacterial partner alters the metabolism, cell membrane composition but not the exoskeleton of the fungus.

Discussion

Mucoromycotina display a great diversity of genome sizes and genome complexity [86,87][88–90]. Previously sequenced *Umbelopsis* genomes are very compact with 21-23 Mb and harbour around 9,100 genes (9,081 *U. isabellina* NBRC 7884 [91], *U. isabellina* AD026: 9,193 genes [92] and *U. ramanniana* AG 9,931 genes [92]). Known Endogonales genomes vary between 96 Mb for *Endogone* sp. and 240 Mb for *Jimgerdemannia flammicorona* and harbour from ca. 9,500 to 16,800 genes [25]. *Mucor* and *Rhizopus* species, depending on the number of whole-genome duplications (WGD), have genomes ranging from 25 Mb up to 55 Mb and up to 17,000 genes [54,92,93].

In the present study, we aimed to point several differences between Umbelopsidales and Mucorales taxa. Besides limited genome size, *Umbelopsis* fungi have been demonstrated to be often associated with bacteria and to produce unsaturated fatty acids [70]. Our results build on top of these observations. Sequenced *Umbelopsis* taxa have compact genomes, almost devoid of repetitive sequences, and contain a moderate number of genes. They have a relatively high number of metabolism-related enzymes, especially glycohydrolases compared to Mucorales, but fewer peptidases. Moreover, we showed that all analysed Mucoromycotina display comparable enzymatic capabilities tested on BIOLOG FF microplates and predicted from genomic data regardless of their genome size. Mucorales and Umbelopsidales representatives are mostly soil-inhabiting saprotrophs and require a broad spectrum of secreted enzymes which is reflected in high count of encoded peptidases, glycohydrolases and transporters. Nevertheless, they differ in

the ability to use particular carbon sources, produce specific proteases, and have different ratios of different classes of phospholipids.

On one hand, *Umbelopsis* spp. in general have fewer secreted peptidases like pepsins and sublitisins than Mucorales. On the other, they encode several families of metalloproteases and cysteine proteases with unknown functions which are absent from Mucorales. The compact genomes of Umbelopsidales are richer in glycohydrolases and CaZymes. In consequence, they have a tendency to use carbohydrates efficiently and grow fast on these carbon sources. When compared to Mucorales, they produce higher amounts of 18:1 fatty acid and have more fucose in their cell wall. However, the meaning of these findings remains to be understood. Despite sharing an ecological niche, these two orders differ in genome size, associated bacteria and degrading capabilities but, nonetheless, share clear synapomorphic traits. Our experimental results support the validity of 18:3 lipids as a chemotaxonomic marker of Mucoromycotina and fucose as a specific component of their cell wall [4,85].

Previous genomic studies covered diverse Mucorales representatives [41,86,87] whereas Umbelopsidales genomes were published in the form of brief genomic reports [91] without a description of the genomic content. In this study, we aimed to fill this knowledge gap by bringing together genomic analyses with phenotype and biochemical studies, especially in the context of how these fungi function in their environment. Vast carbohydrate related enzyme repertoire observed in Umbelopsidales can be related to their ecology. According to the GlobalFungi database [94] the representatives of the genus *Umbelopsis* are present in 6155 out of 20009 global amplicon samples (ca 30%). They are detected mainly in Europe, North America, and Australia, most often in soil, root, and shoots probes from forest or grassland biomes. The representatives of this genus are well-known late wood colonisers, that probably feed on the substrates which were decomposed by other organisms able to degrade complex substrates like cellulose [95]. However, *Umbelopsis* representatives are also often isolated from living plant

material or forest soil [96] and are considered to be plant growth-promoting organisms and root endophytes [97,98]. They can alter plant metabolism leading to the enhanced production of complex metabolites which are not produced without the endophyte [99]. *U. isabellina* and *U. vinacea* described in this study also encoded metabolic clusters including terpenoid clusters.

The representatives of Umbelopsidales were also shown to represent relatively high resistance to some heavy metals like Zn, Mn, Ni or Pb [100] and xenobiotics, such as herbicides [101]. The resistance may be correlated with the presence of numerous ABC transporters encoded in all analysed genomes. Moreover, these genomes have a single arsenite resistance protein homologous to *Absidia repens* BCR42DRAFT_142507 and consisting of ARS2 (PF04959), DUF4187 (PF13821), RRM_1 (PF00076), SERRATE_Ars2_N (PF12066). Such domain architecture is conserved from chytrids to Entomophtoromycotina and Glomeromycotina. Interestingly, the genomes had neither Dikarya-type nor animal metallothioneins.

Umbelopsidales have also been detected in the deep-sea sediments from Magellan seamounts constituting 3.8% of all OTUs [102]. Although this finding is surprising as *Umbelopsis* representatives are well known terrestrial organisms, their presence may be explained by association with plant material. Interestingly, 0.85% of all amplicon samples in which *Umbelopsis* spp. were detected according to GlobalFungi database (Větrovský et al. 2020), are also originating from marine biome. The understanding of this pattern needs further research.

Some representatives of Umbelopsidales are recently considered as effective single cell oils (SCOs) producers as they are capable of producing high amounts of lipids (75% to 84% in dry cell weight; w/w), including polyunsaturated fatty acids (PUFAs). These substances of high dietary and pharmaceutical importance are also considered as precursors for the synthesis of lipid-based biofuels. Although several Mucoromycota representatives were reported to synthesize PUFAs, *U. isabellina* cultivated on glucose has presented exceptionally high lipid production (comparable to

the highest values achieved for genetically engineered SCO-producing bacterial strains) [73]. Analysed strains displayed a typical composition of PUFAs with a predominance of oleic acid (18:1), and higher levels of gamma-linolenic acid (18:3) in Mucorales compared to Umbelopsidales. Surprisingly, *Thamnidium* mycelium showed a relatively high saturation index possibly due to interaction with *Paenibacillus*. Bacteria produce more saturated fatty acids during biofilm formation yet it remains to be elucidated if the bacteria-fungus interaction has a similar effect on fatty acid composition in both partners [103]. *Thamnidium* genome contains a high number of lipid metabolism-related genes and there is no clear explanation for the high level of fatty acid saturation. This phenotype was particularly unexpected since *T. elegans* has been used for biotechnological production of gamma-linoleic acid on diverse substrates especially in low temperature [104–106]. It may be hypothesized that an extended set of lipid-processing enzymes in *Thamnidium* is required in order to balance the bacterial impact on lipid homeostasis and shift it back towards fungal characteristics.

Interestingly, several representatives of Umbelopsidales have recently been shown to be colonized by EHB from Burkholderiaceae. Moreover, the metagenomic analysis of another oleaginous fungus - *Mortierella elongata* and its endosymbiont *Mycoavidus cysteinexigens* showed that bacteria alters the metabolism of the fatty acids of the host. Endosymbiont was shown not only to cause declines in the storage of carbohydrates, organic acids and nitrogenous metabolites but also to be involved in the catabolism of fungal fatty acids and changes volatile compounds profiles of the fungus [32]. Among six sequenced Mucoromycotina, we found numerous bacterial reads which assembled into a complete genome only in *Thamnidium elegans*. The presence of an associated bacteria is reflected in, among else, DAG/TAG lipids composition and utilization of carbohydrates which are accessible exclusively to Dikarya and bacteria. This phenomenon together with a limited set of carbohydrate-processing enzymes present in

Thamnidium excludes the possibility of bacterial contamination and supports the hypothesis of intimate interaction with detected *Paenibacillus*.

Most of the initial reports on intracellular fungal bacteria were based on microscopic observation of uncultivable bacteria inside fungal hyphae [107]. Nowadays, endohyphal bacteria (EHB) can be efficiently identified using genome sequencing methods. The traces of bacterial presence have been found in other published genomes of basal fungi [108]. Identification of EHB can expand our scarce knowledge on the frequency and host range of fungal-bacteria interactions. Our finding of *Paenibacillus* sp. associated with *Thamniudium elegans* is in line with this trend. Diverse *Bacillus* bacteria were found living with truffles [109,110]. *Paenibacillus* has been reported from *Laccaria bicolor* [111], *Sebacina vermifera [112]*, and is known to produce pre-symbiotic and symbiotic interactions with *Glomus [113]*. The genus *Paenibacillus* was erected from *Bacillus* by Ash et al. [114]; it belongs to the family Paenibacillaceae and comprises 253 highly variable species (https://lpsn.dsmz.de/genus/paenibacillus). The representatives of this genus were isolated from a wide range of sources of plant and animal origin. *Paenibacillus* tend to occupy a similar niche to Mucoromycotina moulds as they inhabit soil and dung. The best-studied species - *P. larvae* is known to cause lethal disease of honeybees. However, some other species are known for their plant growth promotion capacities (via siderophores or phytohormones synthesis), others produce a variety of antimicrobials and insecticides. Bacteria from this genus were also shown to produce a plethora of enzymes, like amylases, cellulases, hemicellulases, lipases, pectinases, oxygenases or dehydrogenases [115]. The proteome of *P. larvae* includes a wide range of virulence factors including proteases and toxins [116]. *Paenibacillus validus* stimulated the growth of *Glomus intraradices* [117] and *P. vortex* facilitates dispersal of *Aspergillus fumigatus* [118]. Although the representatives of *Paenibacillus* are known to promote plant growth and secrete several antimicrobial compounds, the endohyphal strain from our study did not encode any antibiotic compounds except for rifamycin-inactivating phosphotransferase. Rather, it provided

31

multiple enzymes that significantly expanded the digestive capabilities of *Thamnidium* while reducing its genome size. Observed genome shrinking cannot be explained solely by the fragmented assembly of *Thamnidium* (its incompleteness is estimated at approximately 5%) because the differences in CaZyme and peptidase abundance from remaining isolates are far greater.

One of the major differences between *Thamnidium* and the remaining isolates is in the lipid composition potentially contributing to cell membrane stiffness. EHB are known to influence host lipid production and bias the TAG/PE ratio [33]. *Thamnidium* had the lowest TAG/PE ratio among tested isolates which might be a sign of symbiotic interaction with *Paenibacillus*, yet the noted ratio was still far from the 1:1 "symbiotic equilibrium" described in Lastovetsky's report [33]. It is not known whether the TAG/PE values estimated for symbiosis between *Rhizopus microsporus* and *Mycetohabitans* are valid also for other models. It is an open question of how intimate and stable is the interaction between *T. elegans* and *Paenibacillus* sp. The altered traits of *Thamnidium elegans* compared to remaining isolates regardless of their evolutionary distances could be explained by the presence of an associated bacterium classified to *Paenibacillus*. In contrast to lipids, the cell wall carbohydrate composition of *T.elegans* remained unchanged. What we observed is that *Thamnidium* has several metabolic parameters altered but its morphology remained unchanged compared to strains without detectable bacterial partners.

The ancestors of extant Mucoromycotina were present among the first land colonizers and had the ability to access decomposing material. Genome sequencing and phenotyping of Mucorales and Umbelopsidales enabled us to look at the differences of these two old lineages within Mucoromycotina. There are several differences, particularly in their carbon source preferences and encoded carbohydrate repertoire, which hints at subtle niche differentiation. Importantly, predicted digestive capabilities are in line with experimental validation. Early diverging Mucoromycotina possess features characteristic of fungi including ergosterol present in the

membranes and a cell wall made of chitin and beta-glucan. Additionally, all studied Mucoromycotina representatives produce 18:3 gamma-linoleic acid and encrust their cell wall with fucose, both of which traits can be a handy discriminant for marking their presence in environmental samples.

Methods

Isolates

Six non-pathogenic, common and soil-borne representatives of Mucorales and Umbelopsidales were chosen for sequencing. Two of them represented the *Umbelopsis* genus, isolated from forest soil in Warsaw (Poland). Other two taxa (i.e. *M. circinatus* and *M. plumbeus*) were also soil-derived isolates from both Americas but representing Mucorales order. Finally, two Mucorales representatives that are well known for their proteolytic activity (i.e. *Mucor saturninus* and *Thamnidium elegans*) were also selected (Table 2). Related organisms were tested for biotechnological usage and reference genomic information for some of them (e.g. *U. isabellina*) was already available. The species-level identification of all isolates was confirmed by sequencing of ITS rDNA fragments (according to the protocol proposed by Walther and co-workers [119]) prior to genome sequencing.

Table 2. Used strains

Species	Strain number	Country of isolation	Substrate
Umbelopsis isabellina	WA0000067209	Poland	forest soil
Umbelopsis vinacea	WA0000051536	Poland (Bielany forest)	forest soil
Mucor circinatus	CBS 142.35	Brazil	soil
Mucor plumbeus	CBS 226.32	Canada	forest soil

Mucor saturninus	WA0000017839	Poland (Warsaw)	cheese
Thamnidium elegans	WA0000018081	Poland (Truskaw)	horse dung

Phenotypic microarray plates

FF phenotypic microarray plates (Biolog Inc., USA) were used to test the capacity of 6 strains to grow on 95 different carbon sources. Carbon sources were grouped into guilds according to Preston-Mafham et al. [69]. All fungal strains were cultured on Potato Dextrose Agar for 7 days and further their swabbed spores were suspended in FF inoculation fluid (deficient amount of carbon sources) to produce a final optical density of 0.036 A at 590 nm. Spores' suspensions were then inoculated on FF microplates and incubated in the aerobic SpectrostarNano universal plate reader (BMR Labtech, Germany) for 96 h at 20 °C. The analysis of each strain was done in three replicates. The metabolic activity was measured kinetically by determining the colorimetric reduction of a tetrazolium dye. Colorimetric values for wells containing carbon substrates were blanked against the control well. The result was considered positive when a difference between the metabolic activities of the first and last day of incubation was observed in all three repetitions. The mean values and standard deviations of AUC (area under the curve) were calculated for each strain and each guild of carbon sources. The metabolic activity of each species on a particular substrate was represented as a heatmap of log-transformed mean AUC values. All analyses were performed in RStudio using packages ggplot2 and vegan v2.4.2).

Lipids analysis

Extraction. Lipids from fungal cultures of the stationary phase of growth were extracted according to the method proposed by Folch et al. [120] with some modifications. The fungal biomass was filtered and 0.1 mg was transferred into Eppendorf tubes containing glass beads, 0.66 mL of chloroform and 0.33 mL of methanol. The homogenization process using a ball mill (FastPrep)

was carried out for 1 min. The mixture was extracted for 2 min. In order to facilitate the separation of two layers, 0.2 mL of 0.9 % saline was added. The lower layer was collected and evaporated.

Phospholipid determination. The polar lipids were measured using an Agilent 1200 HPLC system (Santa Clara, CA, USA) and a 4500 Q-TRAP mass spectrometer (Sciex, Framingham, MA, USA) with an ESI source. For the reversed-phase chromatographic analysis, 10 µL of the lipid extract was injected onto a Kinetex C18 column (50 mm × 2.1 mm, particle size: 5 µm; Phenomenex, Torrance, CA, USA). The mobile phase consisted of 5-mM ammonium formate in water (A) and 5-mM ammonium formate in methanol (B). The solvent gradient was initiated at 70% B, increased to 95% B over 1.25 min, and maintained at 95% B for 6 min before returning to the initial solvent composition over 3 min. The column temperature was maintained at 40 °C, and the flow rate was 500 µL min−1. The instrumental settings of mass spectrometer were as follows: spray voltage −4500 V, curtain gas (CUR) 25, nebulizer gas (GS1) 50, turbo gas (GS2) 60, and ion source temperature of 600 °C. The data analysis was performed with the Analyst™ v1.6.2 software (Sciex, Framingham, MA, USA).

Two approaches were applied to identify PLs: targeted and untargeted. The untargeted approach was performed with the precursor ion scanning (precursor for m/z 153) survey scan, triggering the EPI experiments. On the basis of the untargeted analysis, a comprehensive list of the multiple reaction monitoring (MRM) transitions was generated.

Acylglycerols. DAGs and TAGs analysis was undertaken by liquid chromatography coupled to mass spectrometry (LC-MS) with electrospray ionization (ESI) on an QTRAP 4500 (Sciex). A Kinetex C18 column (see phospholipids determination) and mobile phases consisting of water (A), a mixture of acetonitrile:isopropanol (5:2, v/v) (B) and 5 mM ammonium formate with 0.1 %

formic acid were used. The solvent gradient was initiated at 35% B, increased to 100% B over 4 min, and maintained at 100% B for 11 min before returning to the initial solvent composition over 2 min. The column temperature was maintained at 40 °C, and the flow rate was 600 µL min−1.

The QTRAP 4500 was operated with positive ionization at an electrospray voltage of 5500 V and a targeted multiple reaction monitoring (MRM) approach containing transitions for known precursor/product mass-to-charge ratio. Under these conditions, the TAG ionize as ammonium adducts.

Fatty acid analysis. A lipid sample was diluted in 1.5 mL of methanol and transferred to a screw-capped glass test tube. To the lipid solution, 0.2 mL of toluene and 0.3 mL of the 8.0% HCl solution were added [121]. The tube was vortexed and then incubated at 45 °C overnight. After cooling to room temperature, 1 mL of hexane and 1 mL of water were added for the extraction of fatty acid methyl esters (FAMEs). The tube was vortexed, and then, 0.3 mL of the hexane layer was moved to the chromatographic vial. 1.6 µL of the extract samples were analyzed using gas chromatography.

A FAMEs analysis was performed with an Agilent Model 7890 gas chromatograph, equipped with a 5975C mass detector. The separation was carried out in the capillary column HP 5 MS methyl polysiloxane (30 m × 0.25 mm i.d. × 0.25 mm ft). The column temperature was maintained at 60 °C for 3 min, then increased to 212 °C at the rate of 6 °C min−1, followed by an increase to 245 °C at the rate of 2 °C min−1, and finally to 280 °C at the rate of 20 °C min−1. The column

temperature was maintained at 280 °C for 10 min. Helium was used as the carrier gas at the flow rate of 1 ml min−1. The injection port temperature was 250 °C. The split injection was employed.

Fungal fatty acids were identified by comparison with the retention times of the authentic standards (Sigma, Supelco) and the results were expressed as a percentage of the total amount of fatty acids.

Sterol analysis was undertaken using the QTRAP 3200 (Sciex) mass spectrometer connected to a 1200 series HPLC system. A Kinetex C18 column (see phospholipids determination) was used. The solvents were: water and methanol, both containing 5mM ammonium formate. Analytes were eluted with the following gradient: 40% solvent B from 0 to 1 minute, 100% solvent B from 1 to 4 minutes, 40% solvent B from 4.0 to 4.1 minutes, 40% solvent B from 4.1 to 6 minutes with the flow rate 0.8 ml min^{-1}. The QTRAP instrument was set to the positive ion mode, with the atmospheric pressure chemical ionization (APCI) temperature of 550°C.

Cell wall preparation and determination of cell wall carbohydrates

Fungi were cultivated in PDB medium, washed with 10 mM Tris/HCl, pH 7.5, suspended in the same buffer, disintegrated with 0.5 mm glass beads in the presence of a protease inhibitor cocktail (Sigma-Aldrich) and centrifuged at 1500 x g for 10 min. The resulting pellet containing cell walls was washed with ice-cold 1 M NaCl until the disappearance of absorbance at 260-280 nm [122]. The lyophilized cell wall was hydrolyzed o/n in 4 M trifluoroacetic acid (TFA) at 100° C. After cooling on ice, samples were centrifuged at 17 000 x g for 5 min at 4°C. The supernatant was dried under N2 and washed twice with pure methanol. After removing methanol with N2, the pellet was resuspended in Mili Q water and purified on a Millipore Filter Device (0.45 µm pores) by centrifugation at 16 000 x g for 4 min. Samples were stored at -20°C. Monosaccharides were

37

determined by high-performance anion-exchange chromatography using a Dionex ICS-3000 Ion Chromatography System with a Carbo Pac PA10 analytical column. Neutral sugars were eluted with 18 mM NaOH at 0.25 ml/min [123].

Culture conditions and DNA extraction

All fungal strains were cultured on 4% Potato Dextrose Agar for 7 days at 20ºC. Total genomic DNA was extracted from 30 mg of fresh mycelium following a CTAB-based chloroform extraction protocol [124] and cleaned-up following caesium chloride density gradient centrifugation method [125]. DNA quality and concentration were estimated by 1% agarose gel electrophoresis and NanoDrop® (Thermo Fisher Scientific). Quantification of purified DNA was performed on a Qubit Fluorometer. The identity of all strains was confirmed by the preliminary sequencing of the internal transcribed spacer (ITS rDNA) region and with standard morphological identification procedures.

Sequencing

DNA was sequenced at the High Throughput Sequencing Facility of UNC, Chapel Hill, NC, USA.

Whole-genome sequencing was accomplished using a hybrid approach, combining Illumina short-read data with PacBio long-read data.

Total cellular DNA was sheared using a Covaris E220 sonicator to achieve fragments with an average size of 500 bp. Then libraries were prepared using the Kapa Hyper kit. Libraries were size selected for insert fragments around 500 base pairs using Pippin Prep automatic DNA size selection system (Sage Science). Libraries were analyzed and quantified using a LabChip GX automated electrophoresis system (Caliper) and pooled. The pools were sequenced on Illumina MiSeq sequencer paired-end sequencing (2×300 cycles) to obtain longer reads and on Illumina HiSeq sequencer 2500 (2x150 cycles) to obtain more coverage.

For the PacBio RSII data, ten microgram aliquots of genomic DNA were sheared in a Covaris g-

TUBE to a target fragment size of 20 kb using the shearing conditions provided in the Covaris g-TUBE user manual. The protocol for preparing a 20 kb library (Pacific Biosciences Procedure and Checklist- 20kb Template Preparation Using Blue Pippin™ Size-Selection system) was subsequently followed, using 5 µg of purified, sheared DNA as starting material. Template concentration was calculated using the Qubit fluorometer and the average size was determined by BioAnalyzer trace analysis and served as input to the Annealing & Binding Calculator v.2.1.0.2 (Pacific Biosciences) to prepare SMRTbell-template annealing and polymerase-template binding reactions, as well as the final dilution of the polymerase-bound template complex for sample plate loading and spike-in of control DNA. The PacBio reads were filtered to a minimum read length of 100 bp and a minimum read quality score of 0.85.

Assembly

PacBio reads were processed with SMRT Analysis Software. Spades (v.3.10.1) [126] was chosen as a slow but high-quality assembler [127]. An initial assembly of Illumina reads alone with Abyss [128] produced very fragmented and incomplete assemblies. Quast (v.4.6.1) was used to select the best assembly [129]. The completeness of genome assembly and gene annotation was assessed using BUSCO (v.3b) [38], using BUSCO's fungal dataset universal single-copy orthologs from OrthoDB9 as reference.

BlobTools2 [130] was used to partition the scaffolds, on the basis of read coverage, G/C content, and taxonomic affiliation. The bacterial and fungal reads were assembled separately.

Gene calling

MAKER (v.2.31.8) [131] was used as the main annotation pipeline for gene prediction and annotation. As part of MAKER, AUGUSTUS (3.3) [132], and GeneMark-ES fungal version (v.4.21) were used as *ab initio* gene predictors. The *ab initio* was trained for a couple of times within MAKER based on sequence similarity search via BLAST and sequence alignment via Exonerate

(v.2.2) [133] with default parameter settings, using 100,082 EST and 273,458 protein sequences of Mucoromycotina obtained from NCBI in January 2018, together with SwissProt reference proteins (downloaded 18th January 2018). Repeats were predicted with RepeatMasker (v.4.0.7) [134] together with Tandem Repeat Finder (TRF) (v.4.04) [135] and RMBlast (NCBI blast package of RepeatMasker), based on Repbase database 2017 edition [136] as described in [39]. Gene calling in Paenibacillus was performed with Prokka [137], an antimicrobial component search was done in Abricate [138]. A whole genome-based taxonomic analysis was performed in Type (Strain) Genome Server (TYGS) [139] on 2020-06-01.

Protein coding gene annotation

Predicted protein-coding genes were mapped on Cazy with DBCan [140] modified like in [141], MEROPS protease evalue 1E-10 [142], PFAM evalue 0.001 [143], INTERPROScan 5.28-67 (v[144]), CDD evalue 1E-5 [145], AntiSMASH 4.0 (--transatpks_da --clusterblast --subclusterblast --knownclusterblast --smcogs --cassis --borderpredict --full-hmmer) [146], TCDB evalue 1E-5 [147]. RNAi components were filtered with an evalue of 1E-10. KEGG mapping was performed with kofamscan 1.3.0 with default settings [148].

Phylogenetic analysis of bacteria found in the genome of Thamnidium

16S rRNA sequence was extracted from the bacterial reads found in Thamnidium using blast. Then it was combined with publicly available 16S sequences of several species and strains of *Paenibacillus* and *Bacillus* (GB accession numbers can be found on Fig. 1). Sequences were then aligned using MAFFT [149] and trimmed using trimal -automated1 [150]. Then the best evolution model was detected using modeltest-ng across all evolutionary models [151] (TrN+I+G4 model selected based on AIC, BIC and AICc criteria) and phylogenetic tree was calculated using raxml-ng [152] with 1000 bootstrap replicates. The tree was then rooted using four *Bacillus* sequences.

Acknowledgements

Data Availability Statement

All of the genomic data have been deposited in the NCBI BioProject PRJNA668042. Raw reads for all fungal strains are available in the SRA database under accession numbers SRR12875449-SRR12875464. Assemblies and annotations are deposited under accession numbers XXXXXXX-XXXXXX. Bacterial reads, assembly, and annotation can be found under accession numbers SRAXXXX, XXXXX, and XXXX, respectively.

The sequence analyses are based on publicly available sequences from NCBI and UNIPROT.

Competing interests

The authors declare that they have no competing interests.

Author Contributions

A.M. and J.P. designed the study, A.M. performed assembly, annotation and sequence analyses, J.S.K, P.B., A.O., K.S., J.P. and A.M. interpreted the data and drafted the manuscript, K.S., J.P. and A.M. wrote the manuscript, E.M. and P.M. performed whole-genome sequencing, O.S., U.Z.

41

and S.P. performed experimental procedures and prepared the samples, T.A-P and K.SA. performed carbon source usage experiments U.P-L and J.S.K. analysed cell wall composition, P.B. analysed lipid composition, J.P. analysed carbon source usage.

Supplementary Materials

Supplementary Figure S1. Tree inferred with FastME 2.1.6.1 [153] from Genome BLAST Distance Phylogeny approach distances calculated from genome sequences of Paenibacillus isolates [154]. The numbers above branches are pseudo-bootstrap support values > 60 % from 100 replications, with average branch support of 47.8 %. The tree was rooted at the midpoint by the program. Pae_T is the *Paenibacillus* sequenced as associated with *Thamnidium elegans* groups with *P. illinoisensis* strains.

Supplementary Tables

Supplementary Table S1 Repetitive elements in analysed genomes

Supplementary Table S2 RNAi pathway components in the analysed genomes

Supplementary Table S3 Merops families of peptidases in analysed genomes

Supplementary Table S4 CaZyme families in analysed genomes

Supplementary Table S5 TCDB families of transporters in analysed genomes

Supplementary Table S6 Antismash metabolic clusters in analysed genomes

Supplementary Table S7 Carbon source usage of the isolates

Supplementary Table S8 Total lipid and fatty acid composition of the isolates

Supplementary Table S9 Phospholipid composition of the isolates

Supplementary Table S10 Diacyglicerols and triacyglicerols content in the analysed isolates

Supplementary Table S11 TYGS results for *Paenibacillus* sp. genome listing DDH scores for related organisms.

Literature

1. Spatafora JW, Chang Y, Benny GL, Lazarus K, Smith ME, Berbee ML, et al. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia. 2016;108: 1028–1046.

2. Desirò A, Rimington WR, Jacob A, Pol NV, Smith ME, Trappe JM, et al. Multigene phylogeny of Endogonales, an early diverging lineage of fungi associated with plants. IMA Fungus. 2017;8: 245–257.

3. Richards TA, Leonard G, Wideman JG. What Defines the "Kingdom" Fungi? Microbiol Spectr. 2017;5. doi:10.1128/microbiolspec.FUNK-0044-2017

4. Mélida H, Sain D, Stajich JE, Bulone V. Deciphering the uniqueness of Mucoromycotina cell walls by combining biochemical and phylogenomic approaches. Environ Microbiol. 2015;17: 1649–1662.

5. Bartnicki-Garcia S. Cell Wall Chemistry, Morphogenesis, and Taxonomy of Fungi. Annu Rev Microbiol. 1968;22: 87–108.

6. Field KJ, Rimington WR, Bidartondo MI, Allinson KE, Beerling DJ, Cameron DD, et al. First evidence of mutualism between ancient plant lineages (Haplomitriopsida liverworts) and Mucoromycotina fungi and its response to simulated Palaeozoic changes in atmospheric CO_2. New Phytol. 2015;205: 743–756.

7. Bidartondo MI, Read DJ, Trappe JM, Merckx V, Ligrone R, Duckett JG. The dawn of symbiosis between plants and fungi. Biol Lett. 2011;7: 574–577.

8. Partida-Martinez LP, Hertweck C. Pathogenic fungus harbours endosymbiotic bacteria for toxin production. Nature. 2005;437: 884–888.

9. Ibrahim AS, Spellberg B. Zygomycetes as Agents of Infectious Disease in Humans. Molecular Principles of Fungal Pathogenesis. pp. 428–440.

10. Richardson MD, Rautemaa-Richardson R. Biotic Environments Supporting the Persistence of Clinically Relevant Mucormycetes. J Fungi (Basel). 2019;6. doi:10.3390/jof6010004

11.   Morin-Sardin S, Nodet P, Coton E, Jany J-L. Mucor: A Janus-faced fungal genus with human health impact and industrial applications. Fungal Biology Reviews. 2017. pp. 12–32. doi:10.1016/j.fbr.2016.11.002

12.   Nout MJR, Robert Nout MJ, Aidoo KE. Asian Fungal Fermented Food. Industrial Applications. 2010. pp. 29–58.

13.   Hermet A, Méheust D, Mounier J, Barbier G, Jany J-L. Molecular systematics in the genus Mucor with special regards to species encountered in cheese. Fungal Biol. 2012;116: 692–705.

14.   Rivaldi JD, Carvalho AKF, da Conceição LRV, de Castro HF. Assessing the potential of fatty acids produced by filamentous fungi as feedstock for biodiesel production. Preparative Biochemistry and Biotechnology. 2017. pp. 970–976. doi:10.1080/10826068.2017.1365246

15.   Roukas T, Varzakakou M, Kotzekidou P. From cheese whey to carotenes by Blakeslea trispora in a bubble column reactor. Appl Biochem Biotechnol. 2015;175: 182–193.

16.   Voigt K, Wolf T, Ochsenreiter K, Nagy G, Kaerger K, Shelest E, et al. 15 Genetic and Metabolic Aspects of Primary and Secondary Metabolism of the Zygomycetes. Biochemistry and Molecular Biology. 2016. pp. 361–385.

17.   Satari B, Karimi K. Mucoralean fungi for sustainable production of bioethanol and biologically active molecules. Appl Microbiol Biotechnol. 2018;102: 1097–1117.

18.   Meghwanshi GK, Vashishtha A. Biotechnology of Fungal Lipases. Fungi and their Role in Sustainable Development: Current Perspectives. 2018. pp. 383–411. doi:10.1007/978-981-13-0393-7_22

19.   Silva M da, da Silva M, Manfio GP, Canhos VP. Characterization of selected strains of mucorales using fatty acid profiles. Rev Microbiol. 1998;29: 276–281.

20.   Mukherjee KD, Kiewitt I. Formation of .gamma.-linolenic acid in the higher plant evening primrose (Oenothera biennis L.). Journal of Agricultural and Food Chemistry. 1987. pp. 1009–1012. doi:10.1021/jf00078a036

21.   Nichols BW, Wood BJB. The occurrence and biosynthesis of gamma-linolenic acid in a blue-green alga,Spirulina platensis. Lipids. 1968. pp. 46–50. doi:10.1007/bf02530968

22.   Cornely OA, Alastruey-Izquierdo A, Arenz D, Chen SCA, Dannaoui E, Hochhegger B, et al. Global guideline for the diagnosis and management of mucormycosis: an initiative of the European Confederation of Medical Mycology in cooperation with the Mycoses Study Group Education and Research Consortium. The Lancet Infectious Diseases. 2019. pp. e405–e421. doi:10.1016/s1473-3099(19)30312-3

23.   Neher DA, Weicht TR, Bates ST, Leff JW, Fierer N. Changes in bacterial and fungal communities across compost recipes, preparation methods, and composting times. PLoS One. 2013;8: e79512.

24.   Maheshwari R, Bharadwaj G, Bhat MK. Thermophilic Fungi: Their Physiology and Enzymes. Microbiology and Molecular Biology Reviews. 2000. pp. 461–488. doi:10.1128/mmbr.64.3.461-488.2000

25.   Chang Y, Desirò A, Na H, Sandor L, Lipzen A, Clum A, et al. Phylogenomics of Endogonaceae and evolution of mycorrhizas within Mucoromycota. New Phytol. 2019;222: 511–525.

26.   Desirò A, Salvioli A, Ngonkeu EL, Mondo SJ, Epis S, Faccio A, et al. Detection of a novel intracellular microbiome hosted in arbuscular mycorrhizal fungi. ISME J. 2014;8: 257–270.

27.   Desirò A, Hao Z, Liber JA, Benucci GMN, Lowry D, Roberson R, et al. Mycoplasma-related endobacteria within Mortierellomycotina fungi: diversity, distribution and functional insights into their lifestyle. ISME J. 2018;12: 1743–1757.

28.   Bonfante P, Venice F. Mucoromycota: going to the roots of plant-interacting fungi. Fungal Biology Reviews. 2020. pp. 100–113. doi:10.1016/j.fbr.2019.12.003

29.   Itabangi H, Sephton-Clark PCS, Zhou X, Insua I, Probert M, Correia J, et al. A bacterial endosymbiont enables fungal immune evasion during fatal mucormycete infection. doi:10.1101/584607

30.   Virgianti DP, Natalia D, Nyoman Pugeg Aryantha I. Short Communication: New record of Stenotrophomonas sp. as endosymbiont bacteria in Rhizopus microsporus. Biodiversitas Journal of Biological Diversity. 2020. doi:10.13057/biodiv/d210449

31.   Lumini E, Bianciotto V, Jargeat P, Novero M, Salvioli A, Faccio A, et al. Presymbiotic growth and sporal morphology are affected in the arbuscular mycorrhizal fungus Gigaspora margarita cured of its endobacteria. Cell Microbiol. 2007;9: 1716–1729.

32.   Uehling J, Gryganskyi A, Hameed K, Tschaplinski T, Misztal PK, Wu S, et al. Comparative genomics of Mortierella elongata and its bacterial endosymbiont Mycoavidus cysteinexigens. Environ Microbiol. 2017. doi:10.1111/1462-2920.13669

33.   Lastovetsky OA, Gaspar ML, Mondo SJ, LaButti KM, Sandor L, Grigoriev IV, et al. Lipid metabolic changes in an early divergent fungus govern the establishment of a mutualistic symbiosis with endobacteria. Proc Natl Acad Sci U S A. 2016;113: 15102–15107.

34.   Salvioli A, Ghignone S, Novero M, Navazio L, Venice F, Bagnaresi P, et al. Symbiosis with an endobacterium increases the fitness of a mycorrhizal fungus, raising its bioenergetic potential. ISME J. 2016;10: 130–144.

35.   Mondo SJ, Lastovetsky OA, Gaspar ML, Schwardt NH, Barber CC, Riley R, et al. Bacterial endosymbionts influence host sexuality and reveal reproductive genes of early divergent fungi. Nat Commun. 2017;8: 1843.

36.   Partida-Martinez LP, Monajembashi S, Greulich K-O, Hertweck C. Endosymbiont-dependent host reproduction maintains bacterial-fungal mutualism. Curr Biol. 2007;17: 773–777.

37.   Hanagasaki T, Asato N. Changes in free amino acid content and hardness of beef while dry-aging with : Changes in the quality of beef while dry-aging with. Hanguk Tongmul Chawon Kwahakhoe Chi. 2018;60: 19.

38.   Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31: 3210–3212.

39.   Muszewska A, Steczkiewicz K, Stepniewska-Dziubinska M, Ginalski K. Cut-and-Paste Transposons in Fungi with Diverse Lifestyles. Genome Biol Evol. 2017;9: 3463–3477.

40.   Muszewska A, Steczkiewicz K, Ginalski K. DIRS and Ngaro Retrotransposons in Fungi. PLoS One. 2013;8: e76319.

41.   Lebreton A, Corre E, Jany J-L, Brillet-Guéguen L, Pèrez-Arques C, Garre V, et al. Comparative genomics applied to Mucor species with different lifestyles. BMC Genomics. 2020;21: 135.

42.   Castanera R, Pérez G, López L, Sancho R, Santoyo F, Alfaro M, et al. Highly expressed captured genes and cross-kingdom domains present in Helitrons create novel diversity in Pleurotus ostreatus and other fungi. BMC Genomics. 2014;15: 1071.

43.   Cervantes M, Vila A, Nicolás FE, Moxon S, de Haro JP, Dalmay T, et al. A single argonaute gene participates in exogenous and endogenous RNAi and controls cellular functions in the basal fungus Mucor circinelloides. PLoS One. 2013;8: e69283.

44.   Nicolás FE, Vila A, Moxon S, Cascales MD, Torres-Martínez S, Ruiz-Vázquez RM, et al. The RNAi machinery controls distinct responses to environmental signals in the basal fungus Mucor circinelloides. BMC Genomics. 2015. doi:10.1186/s12864-015-1443-2

45.   Calo S, Nicolás FE, Lee SC, Vila A, Cervantes M, Torres-Martinez S, et al. A non-canonical RNA degradation pathway suppresses RNAi-dependent epimutations in the human fungal pathogen Mucor circinelloides. PLoS Genet. 2017;13: e1006686.

46.   Zhang H-Z, Chambers HF. BlaR1 and MecR1 gene products of Staphylococcus aureus. Handbook of Proteolytic Enzymes. 2004. pp. 769–771. doi:10.1016/b978-0-12-079611-3.50233-0

47.   Tang G-L. Faculty Opinions recommendation of Resistance to nonribosomal peptide antibiotics mediated by D-stereospecific peptidases. Faculty Opinions – Post-Publication Peer Review of the Biomedical Literature. 2018. doi:10.3410/f.732753547.793544905

48.   Dashdorj D, Tripathi VK, Cho S, Kim Y, Hwang I. Dry aging of beef; Review. Hanguk Tongmul Chawon Kwahakhoe Chi. 2016;58: 20.

49.   Navais R, Méndez J, Pérez-Pascual D, Cascales D, Guijarro JA. The yrpAB operon of Yersinia ruckeri encoding two putative U32 peptidases is involved in virulence and induced under microaerobic conditions. Virulence. 2014;5: 619–624.

50.   Kavermann H, Burns BP, Angermuller K, Odenbreit S, Fischer W, Melchers K, et

al. Identification and characterization of Helicobacter pylori genes essential for gastric colonization. J Exp Med. 2003;197: 813–822.

51.   Tallant C, García-Castellanos R, Marrero A, Canals F, Yang Y, Reymond J-L, et al. Activity of ulilysin, an archaeal PAPP-A-related gelatinase and IGFBP protease. Biol Chem. 2007;388: 1243–1253.

52.   Sasagawa Y, Izaki K, Matsubara Y, Suzuki K, Kojima H, Kamio Y. Molecular cloning and sequence analysis of the gene encoding the collagenase from Cytophaga sp. L43-1 strain. Biosci Biotechnol Biochem. 1995;59: 2068–2073.

53.   Maia JC. Hexosamine and cell wall biogenesis in the aquatic fungus Blastocladiella emersonii. FASEB J. 1994;8: 848–853.

54.   Corrochano LM, Kuo A, Marcet-Houben M, Polaino S, Salamov A, Villalobos-Escobedo JM, et al. Expansion of Signal Transduction Pathways in Fungi by Extensive Genome Duplication. Curr Biol. 2016;26: 1577–1584.

55.   Su H, Blain F, Musil RA, Zimmermann JJ, Gu K, Bennett DC. Isolation and expression in Escherichia coli of hepB and hepC, genes coding for the glycosaminoglycan-degrading enzymes heparinase II and heparinase III, respectively, from Flavobacterium heparinum. Appl Environ Microbiol. 1996;62: 2723–2734.

56.   Janesch B, Koerdt A, Messner P, Schäffer C. The S-Layer Homology Domain-Containing Protein SlhA from Paenibacillus alvei CCM 2051T Is Important for Swarming and Biofilm Formation. PLoS ONE. 2013. p. e76566. doi:10.1371/journal.pone.0076566

57.   Baldeweg F, Warncke P, Fischer D, Gressler M. Fungal Biosurfactants from Mortierella alpina. Org Lett. 2019;21: 1444–1448.

58.   Blankenfeldt W, Kuzin AP, Skarina T, Korniyenko Y, Tong L, Bayer P, et al. Structure and function of the phenazine biosynthetic protein PhzF from Pseudomonas fluorescens. Proc Natl Acad Sci U S A. 2004;101: 16431–16436.

59.   Bastias DA, Alejandra Martínez-Ghersa M, Ballaré CL, Gundel PE. Epichloë Fungal Endophytes and Plant Defenses: Not Just Alkaloids. Trends in Plant Science. 2017. pp. 939–948. doi:10.1016/j.tplants.2017.08.005

60.   Orlowska M, Steczkiewicz K, Muszewska A. Utilization of cobalamin is ubiquitous in early-branching fungal phyla. doi:10.1101/2020.10.13.337048

61.   Čertík M, Adamechová Z, Guothová L. Simultaneous enrichment of cereals with polyunsaturated fatty acids and pigments by fungal solid state fermentations. Journal of Biotechnology. 2013. pp. 130–134. doi:10.1016/j.jbiotec.2013.03.016

62.   Avalos J, Prado-Cabrero A, Estrada AF. Neurosporaxanthin production by Neurospora and Fusarium. Methods Mol Biol. 2012;898: 263–274.

63.   Sahadevan Y, Richter-Fecken M, Kaerger K, Voigt K, Boland W. Early and late trisporoids differentially regulate β-carotene production and gene transcript Levels in the mucoralean fungi Blakeslea trispora and Mucor mucedo. Appl Environ Microbiol. 2013;79: 7466–7475.

64.   Lee SC, Heitman J. Sex in the Mucoralean fungi. Mycoses. 2014;57 Suppl 3: 18–24.

65.   Lee SC, Idnurm A. Fungal Sex: The Mucoromycota. Microbiol Spectr. 2017;5. doi:10.1128/microbiolspec.FUNK-0041-2017

66.   Wetzel J, Burmester A, Kolbe M, Wöstemeyer J. The mating-related loci sexM and sexP of the zygomycetous fungus Mucor mucedo and their transcriptional regulation by trisporoid pheromones. Microbiology. 2012;158: 1016–1023.

67.   Idnurm A, Walton FJ, Floyd A, Heitman J. Identification of the sex genes in an early diverged fungus. Nature. 2008. pp. 193–196. doi:10.1038/nature06453

68.   Schulz E, Wetzel J, Burmester A, Ellenberger S, Siegmund L, Wostemeyer J. Sex loci of homothallic and heterothallic Mucorales. Endocytobiosis Cell Res. 2016;27: 39–57.

69.   Preston-Mafham J, Boddy L, Randerson PF. Analysis of microbial community functional diversity using sole-carbon-source utilisation profiles - a critique. FEMS Microbiol Ecol. 2002;42: 1–14.

70.   Fakas S, Makri A, Mavromati M, Tselepi M, Aggelis G. Fatty acid composition in lipid fractions lengthwise the mycelium of Mortierella isabellina and lipid production by solid state fermentation. Bioresour Technol. 2009;100: 6118–6120.

71.   Bernat P, Nykiel-Szymańska J, Stolarek P, Słaba M, Szewczyk R, Różalska S. 2,4-dichlorophenoxyacetic acid-induced oxidative stress: Metabolome and membrane modifications in Umbelopsis isabellina, a herbicide degrader. PLoS One. 2018;13: e0199677.

72.   Demir M, Turhan I, Kucukcetin A, Alpkent Z. Oil production by Mortierella isabellina from whey treated with lactase. Bioresour Technol. 2013;128: 365–369.

73.   Papanikolaou S, Aggelis G. Sources of microbial oils with emphasis to Mortierella (Umbelopsis) isabellina fungus. World J Microbiol Biotechnol. 2019;35: 63.

74.   Wang D, Li M, Wei D, Cai YI, Zhang Y, Xing L. Identification and Functional Characterization of the Delta 6-Fatty Acid Desaturase Gene from Thamnidium elegans. The Journal of Eukaryotic Microbiology. 2007. pp. 110–117. doi:10.1111/j.1550-7408.2006.00136.x

75.   Zhu Y, Zhang B-B. Molecular cloning and functional characterization of a Δ6-fatty acid desaturase gene fromRhizopus oryzae. Journal of Basic Microbiology. 2013. pp. 773–777. doi:10.1002/jobm.201200189

76.   Huang Y-S, Chaudhary S, Thurmond JM, Bobik EG, Yuan L, Chan GM, et al. Cloning of Δ12- and Δ6-desaturases from Mortierella alpina and recombinant production of γ-linolenic acid in Saccharomyces cerevisiae. Lipids. 1999. pp. 649–659. doi:10.1007/s11745-999-0410-8

77.   Darwish E, Testerink C, Khalil M, El-Shihy O, Munnik T. Phospholipid signaling responses in salt-stressed rice leaves. Plant Cell Physiol. 2009;50: 986–997.

78.   Bernat P, Gajewska E, Szewczyk R, Słaba M, Długoński J. Tributyltin (TBT) induces oxidative stress and modifies lipid profile in the filamentous fungus Cunninghamella elegans. Environ Sci Pollut Res Int. 2014;21: 4228–4235.

79.   Renne MF, de Kroon AIPM. The role of phospholipid molecular species in determining the physical properties of yeast membranes. FEBS Lett. 2018;592: 1330–1345.

80.   Gidden J, Denson J, Liyanage R, Ivey DM, Lay JO. Lipid Compositions in Escherichia coli and Bacillus subtilis During Growth as Determined by MALDI-TOF and TOF/TOF Mass Spectrometry. Int J Mass Spectrom. 2009;283: 178–184.

81.   Bernat P, Nesme J, Paraszkiewicz K, Schloter M, Płaza G. Characterization of Extracellular Biosurfactants Expressed by a Pseudomonas putida Strain Isolated from the Interior of Healthy Roots from Sida hermaphrodita Grown in a Heavy Metal Contaminated Soil. Current Microbiology. 2019. pp. 1320–1329. doi:10.1007/s00284-019-01757-x

82.   Bernat P, Paraszkiewicz K, Siewiera P, Moryl M, Płaza G, Chojniak J. Lipid composition in a strain of Bacillus subtilis, a producer of iturin A lipopeptides that are active against uropathogenic bacteria. World Journal of Microbiology and Biotechnology. 2016. doi:10.1007/s11274-016-2126-0

83.   Markgraf DF, Klemm RW, Junker M, Hannibal-Bach HK, Ejsing CS, Rapoport TA. An ER protein functionally couples neutral lipid metabolism on lipid droplets to membrane lipid synthesis in the ER. Cell Rep. 2014;6: 44–55.

84.   Jorasch P, Wolter FP, Zähringer U, Heinz E. A UDP glucosyltransferase from Bacillus subtilis successively transfers up to four glucose residues to 1,2-diacylglycerol: expression of ypfP in Escherichia coli and structural analysis of its reaction products. Mol Microbiol. 1998;29: 419–430.

85.   Lecointe K, Cornu M, Leroy J, Coulon P, Sendid B. Polysaccharides Cell Wall Architecture of Mucorales. Frontiers in Microbiology. 2019. doi:10.3389/fmicb.2019.00469

86.   Tang X, Zhao L, Chen H, Chen YQ, Chen W, Song Y, et al. Complete Genome Sequence of a High Lipid-Producing Strain of Mucor circinelloides WJ11 and Comparative Genome Analysis with a Low Lipid-Producing Strain CBS 277.49. PLoS One. 2015;10: e0137543.

87.   Ma L-J, Ibrahim AS, Skory C, Grabherr MG, Burger G, Butler M, et al. Genomic analysis of the basal lineage fungus Rhizopus oryzae reveals a whole-genome duplication. PLoS Genet. 2009;5: e1000549.

88.   Linde J, Schwartze V, Binder U, Lass-Flörl C, Voigt K, Horn F. De Novo Whole-Genome Sequence and Genome Annotation of Lichtheimia ramosa. Genome Announc. 2014;2. doi:10.1128/genomeA.00888-14

89.   Schwartze VU, Winter S, Shelest E, Marcet-Houben M, Horn F, Wehner S, et al. Gene expansion shapes genome architecture in the human pathogen Lichtheimia corymbifera: an evolutionary genomics analysis in the ancient terrestrial mucorales (Mucoromycotina). PLoS Genet. 2014;10: e1004496.

90.   Chibucos MC, Etienne KA, Orvis J, Lee H, Daugherty S, Lockhart SR, et al. The

genome sequence of four isolates from the family Lichtheimiaceae. Pathog Dis. 2015;73. doi:10.1093/femspd/ftv024

91.   Takeda I, Tamano K, Yamane N, Ishii T, Miura A, Umemura M, et al. Genome Sequence of the Mucoromycotina Fungus Umbelopsis isabellina, an Effective Producer of Lipids. Genome Announc. 2014;2. doi:10.1128/genomeA.00071-14

92.   Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, et al. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res. 2014;42: D26–31.

93.   Gryganskyi AP, Golan J, Dolatabadi S, Mondo S, Robb S, Idnurm A, et al. Phylogenetic and Phylogenomic Definition of Rhizopus Species. G3: Genes|Genomes|Genetics. 2018. pp. 2007–2018. doi:10.1534/g3.118.200235

94.   Větrovský T, Morais D, Kohout P, Lepinay C, Algora C, Awokunle Hollá S, et al. GlobalFungi, a global database of fungal occurrences from high-throughput-sequencing metabarcoding studies. Sci Data. 2020;7: 228.

95.   Richardson M. The ecology of the Zygomycetes and its impact on environmental exposure. Clinical Microbiology and Infection. 2009. pp. 2–9. doi:10.1111/j.1469-0691.2009.02972.x

96.   Sheng Y, Cong J, Lu H, Yang L, Liu Q, Li D, et al. Broad-leaved forest types affect soil fungal community structure and soil organic carbon contents. Microbiologyopen. 2019;8: e874.

97.   Tejesvi MV, Sauvola T, Pirttilä AM, Ruotsalainen AL. Neighboring Deschampsia flexuosa and Trientalis europaea harbor contrasting root fungal endophytic communities. Mycorrhiza. 2013;23: 1–10.

98.   Huang Q, An H, Song H, Mao H, Shen W, Dong J. Diversity and biotransformative potential of endophytic fungi associated with the medicinal plant Kadsura angustifolia. Res Microbiol. 2015;166: 45–55.

99.   Qin D, Wang L, Han M, Wang J, Song H, Yan X, et al. Effects of an Endophytic Fungus Umbelopsis dimorpha on the Secondary Metabolites of Host–Plant Kadsura angustifolia. Frontiers in Microbiology. 2018. doi:10.3389/fmicb.2018.02845

100.   Janicki T, Długoński J, Krupiński M. Detoxification and simultaneous removal of phenolic xenobiotics and heavy metals with endocrine-disrupting activity by the non-ligninolytic fungus Umbelopsis isabellina. J Hazard Mater. 2018;360: 661–669.

101.   Nykiel-Szymańska J, Stolarek P, Bernat P. Elimination and detoxification of 2,4-D by Umbelopsis isabellina with the involvement of cytochrome P450. Environ Sci Pollut Res Int. 2018;25: 2738–2743.

102.   Yang S, Xu W, Gao Y, Chen X, Luo Z-H. Fungal diversity in deep-sea sediments from Magellan seamounts environment of the western Pacific revealed by high-throughput Illumina sequencing. J Microbiol. 2020;58: 841–852.

103.   Dubois-Brissonnet F, Trotier E, Briandet R. The Biofilm Lifestyle Involves an

Increase in Bacterial Membrane Saturated Fatty Acids. Front Microbiol. 2016;7: 1673.

104.	Zikou E, Chatzifragkou A, Koutinas AA, Papanikolaou S. Evaluating glucose and xylose as cosubstrates for lipid accumulation and γ-linolenic acid biosynthesis of Thamnidium elegans. J Appl Microbiol. 2013;114: 1020–1032.

105.	Liu G-Q, Jin X-C. Screening and optimization of microbial lipid production by Thamnidium sp., a novel oleaginous fungus isolated from forest soil. Journal of Biotechnology. 2008. p. S434. doi:10.1016/j.jbiotec.2008.07.1005

106.	Stredansky M, Conti E, Stredanska S, Zanetti F. γ-Linolenic acid production with Thamnidium elegans by solid-state fermentation on apple pomace. Bioresource Technology. 2000. pp. 41–45. doi:10.1016/s0960-8524(99)00132-7

107.	Macdonald RM, Chandler MR. BACTERIUM-LIKE ORGANELLES IN THE VESICULAR-ARBUSCULAR MYCORRHIZAL FUNGUS GLOMUS CALEDONIUS. New Phytologist. 1981. pp. 241–246. doi:10.1111/j.1469-8137.1981.tb07486.x

108.	Naranjo Ortiz MÁ, Others. Non-vertical genomics in fungal evolution. Universitat Pompeu Fabra. 2019.

109.	Barbieri E, Bertini L, Rossi I, Ceccaroli P, Saltarelli R, Guidi C, et al. New evidence for bacterial diversity in the ascoma of the ectomycorrhizal fungus Tuber borchii Vittad. FEMS Microbiol Lett. 2005;247: 23–35.

110.	Perlińska-Lenart U, Piłsyk S, Gryz E, Turło J, Hilszczańska D, Kruszewska JS. Identification of bacteria and fungi inhabiting fruiting bodies of Burgundy truffle (Tuber aestivum Vittad.). Arch Microbiol. 2020;202: 2727–2738.

111.	Bertaux J, Schmid M, Prevost-Boure NC, Churin JL, Hartmann A, Garbaye J, et al. In situ identification of intracellular bacteria related to Paenibacillus spp. in the mycelium of the ectomycorrhizal fungus Laccaria bicolor S238N. Appl Environ Microbiol. 2003;69: 4243–4248.

112.	Sharma M, Schmid M, Rothballer M, Hause G, Zuccaro A, Imani J, et al. Detection and identification of bacteria intimately associated with fungi of the orderSebacinales. Cellular Microbiology. 2008. pp. 2235–2246. doi:10.1111/j.1462-5822.2008.01202.x

113.	Bidondo LF, Silvani V, Colombo R, Pérgola M, Bompadre J, Godeas A. Pre-symbiotic and symbiotic interactions between Glomus intraradices and two Paenibacillus species isolated from AM propagules. In vitro and in vivo assays with soybean (AG043RG) as plant host. Soil Biology and Biochemistry. 2011. pp. 1866–1872. doi:10.1016/j.soilbio.2011.05.004

114.	Ash C, Priest FG, Collins MD. Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Proposal for the creation of a new genus Paenibacillus. Antonie Van Leeuwenhoek. 1993;64: 253–260.

115.	Grady EN, MacDonald J, Liu L, Richman A, Yuan Z-C. Current knowledge and perspectives of Paenibacillus: a review. Microb Cell Fact. 2016;15: 203.

116.   Erban T, Zitek J, Bodrinova M, Talacko P, Bartos M, Hrabak J. Comprehensive proteomic analysis of exoproteins expressed by ERIC I, II, III and IV Paenibacillus larvae genotypes reveals a wide range of virulence factors. Virulence. 2019. pp. 363–375. doi:10.1080/21505594.2019.1603133

117.   Hildebrandt U, Ouziad F, Marner F-J, Bothe H. The bacterium Paenibacillus validus stimulates growth of the arbuscular mycorrhizal fungus Glomus intraradices up to the formation of fertile spores. FEMS Microbiol Lett. 2006;254: 258–267.

118.   Ingham CJ, Kalisman O, Finkelshtein A, Ben-Jacob E. Mutually facilitated dispersal between the nonmotile fungus Aspergillus fumigatus and the swarming bacterium Paenibacillus vortex. Proc Natl Acad Sci U S A. 2011;108: 19731–19736.

119.   Walther G, Pawłowska J, Alastruey-Izquierdo A, Wrzosek M, Rodriguez-Tudela JL, Dolatabadi S, et al. DNA barcoding in Mucorales: an inventory of biodiversity. Persoonia. 2013;30: 11–47.

120.   Folch J, Lees M, Sloane Stanley GH. A simple method for the isolation and purification of total lipides from animal tissues. J Biol Chem. 1957;226: 497–509.

121.   Ichihara K 'ichi, Fukubayashi Y. Preparation of fatty acid methyl esters for gas-liquid chromatography. J Lipid Res. 2010;51: 635–640.

122.   Nemčovič M, Farkaš V. Cell-Wall composition and polysaccharide synthhase activity changes following photoinduction in trichoderma viride. Acta Biologica Hungarica. 2001. pp. 281–288. doi:10.1556/abiol.52.2001.2-3.12

123.   Zdebska E, Kościelak J. A single-sample method for determination of carbohydrate and protein contents glycoprotein bands separated by sodium dodecyl sulfate- polyacrylamide gel electrophoresis. Anal Biochem. 1999;275: 171–179.

124.   Doyle J. DNA Protocols for Plants. Molecular Techniques in Taxonomy. 1991. pp. 283–293. doi:10.1007/978-3-642-83962-7_18

125.   Garber RC, Yoder OC. Isolation of DNA from filamentous fungi and separation into nuclear, mitochondrial, ribosomal, and plasmid components. Anal Biochem. 1983;135: 416–422.

126.   Antipov D, Korobeynikov A, McLean JS, Pevzner PA. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. Bioinformatics. 2015;32: 1009–1015.

127.   Abbas MM, Malluhi QM, Balakrishnan P. Assessment of de novo assemblers for draft genomes: a case study with fungal genomes. BMC Genomics. 2014;15 Suppl 9: S10.

128.   Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. ABySS: A parallel assembler for short read sequence data. Genome Research. 2009. pp. 1117–1123. doi:10.1101/gr.089532.108

129.   Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29: 1072–1075.

130.   Laetsch DR, Blaxter ML. BlobTools: Interrogation of genome assemblies. F1000Research. 2017. p. 1287. doi:10.12688/f1000research.12232.1

131.   Campbell MS, Holt C, Moore B, Yandell M. Genome Annotation and Curation Using MAKER and MAKER-P. Current Protocols in Bioinformatics. 2014. pp. 4.11.1–4.11.39.

132.   Hoff KJ, Stanke M. WebAUGUSTUS--a web service for training AUGUSTUS and predicting genes in eukaryotes. Nucleic Acids Res. 2013;41: W123–8.

133.   Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics. 2005;6: 31.

134.   Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0. 2013--2015. 2015.

135.   Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27: 573–580.

136.   Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob DNA. 2015;6: 11.

137.   Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30: 2068–2069.

138.   Seemann T. abricate. Github. 2019.

139.   Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun. 2019;10: 2182.

140.   Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012;40: W445–51.

141.   Muszewska A, Piłsyk S, Perlińska-Lenart U, Kruszewska JS. Diversity of Cell Wall Related Proteins in Human Pathogenic Fungi. J Fungi (Basel). 2017;4. doi:10.3390/jof4010006

142.   Rawlings ND, Bateman A. How to use the MEROPS database and website to help understand peptidase specificity. Protein Science. 2020. doi:10.1002/pro.3948

143.   El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. Nucleic Acids Res. 2019;47: D427–D432.

144.   Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014. pp. 1236–1240. doi:10.1093/bioinformatics/btu031

145.   Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res. 2020;48: D265–D268.

146.   Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, et al. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res. 2017;45: W36–W41.

147.   Saier MH Jr, Reddy VS, Tsu BV, Ahmed MS, Li C, Moreno-Hagelsieb G. The Transporter Classification Database (TCDB): recent advances. Nucleic Acids Res. 2016;44:

D372–9.

148.   Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. J Mol Biol. 2016;428: 726–731.

149.   Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol. 2013;30: 772–780.

150.   Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25: 1972–1973.

151.   Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. doi:10.1101/612903

152.   Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics. 2019;35: 4453–4455.

153.   Lefort V, Desper R, Gascuel O. FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program: Table 1. Molecular Biology and Evolution. 2015. pp. 2798–2800. doi:10.1093/molbev/msv150

154.   Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics. 2013;14: 60.