# Pan-Genome Analysis Reveals The Molecular Targets of Identification and Virulence Detection in Mucoralean Fungi

**Meijie Zhang**
Chinese Academy of Medical Sciences and Peking Union Medical College Hospital and Institute of Dermatology: Affiliated Hospital for Skin Diseases of Chinese Academy of Medical Sciences

**Wenqi Xu**
Chinese Academy of Medical Sciences and Peking Union Medical College Hospital and Institute of Dermatology: Affiliated Hospital for Skin Diseases of Chinese Academy of Medical Sciences

**Huan Mei**
Chinese Academy of Medical Sciences and Peking Union Medical College Hospital and Institute of Dermatology: Affiliated Hospital for Skin Diseases of Chinese Academy of Medical Sciences

**Naicen Ge**
Chinese Academy of Medical Sciences and Peking Union Medical College Hospital and Institute of Dermatology: Affiliated Hospital for Skin Diseases of Chinese Academy of Medical Sciences

**Ge Song**
Chinese Academy of Medical Sciences and Peking Union Medical College Hospital and Institute of Dermatology: Affiliated Hospital for Skin Diseases of Chinese Academy of Medical Sciences

**Ye Tao**
shanghai biozeron biotechnology Co Ltd

**Weida Liu**
Chinese Academy of Medical Sciences Institute of Dermatology: Affiliated Hospital for Skin Diseases of Chinese Academy of Medical Sciences

**Guanzhao Liang** ( guanzhaoguan@126.com )
Chinese Academy of Medical Sciences and Peking Union Medical College Hospital and Institute of Dermatology: Affiliated Hospital for Skin Diseases of Chinese Academy of Medical Sciences

---

**Research**

**Keywords:** Mucorales, epidemiology, molecular targets, fungal virulence

**Posted Date:** February 28th, 2022

**DOI:** https://doi.org/10.21203/rs.3.rs-1346045/v1

# Abstract

Mucoralean fungi offer various pathogens to cause mucormycosis, especially in immunodeficient patients. Over the past decades, both the morbidity and mortality of mucormycosis have increased rapidly, particularly in developing countries. Nowadays, mucormycosis more often happens in India for the COVID-19 pandemic and its backward diagnostic techniques. Our epidemiologic outcomes show several identifications of Mucoralean fungi are limited to genus, while *Rhizopus* species, *Mucor* species and *Lichtheimia* species have high proportions. To find more molecular targets to make rapid and accurate identifications of *Mucorales* genus and species, Pan-genome analysis and Phylogenetic tree are conducted with four *Mucorales* isolates we sequenced and 43 fungi from NCBI. A few *Mucorales*-specific genes have been found such as STE/STE20 protein kinase, GH36 and sel1 repeat protein. *Mucorales* genus-specific genes are also found in *Lichtheimia* species and *Cunninghamella* species, which covered cellular structure, biochemistry metabolism, molecular processing, and signal transduction. Reported proteins related to the virulence of *Mucorales* species were run with Orthofinder and *112092*, *cotH3*, *gcn4* and *igp1* have shown the potential to be the direct identification as well as the virulence detection of *Mucorales* species. The molecular biological techniques need to be promoted, for which our study provide hypothesis and feasibility analysis.


# 1. Introduction

*Mucorales* can cause mucormycosis in human beings, of which different genus and species often lead to distinct characteristics of infections. As a well-known life-threatening disease, mucormycosis especially occurs in individuals with immunodeficiency [1], diabetes mellitus [2] or major trauma [3]. A review of global epidemiology of mucormycosis had analyzed 851 cases from January 2000 to January 2017 in diverse aspects [4], illustrating a detailed spectrum of causative pathogens. As a result, *Rhizopus arrhizus* (70/447) accounted for the largest proportion, followed by *Lichtheimia corymbifera* (41/447). The review had also pointed out that the disease burden in Europe was higher than that in Asia, which was inconsistent with other studies reporting the incidence of mucormycosis rising much higher in India and China with uncontrolled diabetes mellitus [2, 5–7]. To correct the incidence data error, a recent review studied population-based prevalence reported from various countries, and found the incidence was far underestimated in India and other Asian countries for under-reporting during that period [8].

Since Coronavirus Disease 2019 (COVID-19) burst, multiple case reports have been described co-infected with mucormycosis, most of which are from India, especially in diabetic patients and those overusing corticosteroids, leading to the higher fatality and the worse pandemic outcome [9]. As mentioned in one earlier retrospective analysis of SARS and influenza data from worldwide, *Mucorales* should been called for high attention in this COVID-19 pandemic [10]. While the mucormycosis is more common in developing countries, their laboratory facility is sub-optimal for identification to *Mucorales* genus and species which delays the precise treatment. Despite of available diagnostic technologies proposed as direct or fluorescent microscopy and culture of specimens for morphological identification or DNA

sequencing, promising DNA targets are needed to detect fungi directly in serum and other available body fluids.

Considering the rising attention of proteins associated to *Mucorales* virulence, several DNA targets have been demonstrated potential to detect mucormycosis and to generate novel antifungals against *Mucorales*. Genes regulating the virulence-specific response were mainly revealed through the transcriptomic response of *Mucor circinelloides* (*M. circinelloides*) and validated in vivo [11], including *atf1* (activating transcription factor 1), *atf2* (activating transcription factor 2), *gcn4* (general control nonderepressible 4), *aqp1* (aquaporin 1), *ico1* (isochorismatase 1), *igp1* (immunoglobulin-like protein) and *pps1* (signal peptide-containing proteins), performing important roles in *Mucor*ales survival following phagocytosis, germination inside the phagosome, or the virulence in mice. Besides, the genes *mcplD* (*M. circinelloides* Phospholipase D like protein) and *mcmyo5* (the gene encoding *M. circinelloides* Myosin 5) have been found to maintain full virulence of *M. circinelloides* and validated both in *Galleria mellonella* and mice [12], and the deletion of gene *112092* which codes for a hypothetical extracellular protein of unknown function can result in significant reduction of virulence [13]. To the contrary, deletion of *arf1* (ADP-ribosylation factor) in *M. circinelloides* has been reported increasing the strain virulence [14], *arl1* (*arf*-like proteins) can influence mycelial growth in distinct transcriptional expression patterns [15]. Nagy et al found that the disruption of *hmgR* (3-hydroxy-3-methylglutaryl coenzyme A reductase) genes seem to affect *M. circinelloides* in germination of spores and adaptation to the oxidative stress [16]. As the inhibitors of *hmgR*, *statins* had been reported that they can decrease the virulence of *Rhizopus arrhizus* [17].

Ibrahim et al have published a series of researches about the frequent pathogenic *Mucorales* fungi *Rhizopus arrhizus*. They proposed *cotH3* (the spore coat protein homolog) cell surface proteins [18], the fungal ligands that mediate attachment to *GRP78* during host cell invasion, as well as *fob* (the ferrioxamine binding) cell surface proteins [19], the fungal receptors that mediate attachment to ferrioxamine, show requirements in full virulence of *Rhizopus arrhizus* in vivo. In the current study, we sought to identify the *Mucorales*-specific genes and their respective roles in the diverse pathogenesis of mucormycosis through whole genome sequence, pan-genome analysis and orthofinder.

# 2. Materials And Methods

# 2.1. Search strategies and data analysis

A systematic search comprising keywords and MeSH (Details can be seen in supplementary material Appendix S1.) for causative pathogens was carried out in PUBMED database. Searches were limited to studies involving humans, published in English from 2017 to 2021, of which the final search performed in July 2021. For exclusion, the published cases were removed with conditions of: (i) irrelevant titles or abstracts, (ii) full text unavailable, (iii) non-English articles, (iv) editorial or review, (v) pool identification, (vi) wrong fungi taxonomy.

## 2.2. Fungi isolates collection and DNA sequencing library construction

Clinical isolates *Mucor irregularis* (B50a), *Lichtheimia corymbifera* (B63a), *Mucor hiemalis* (B66h) and *Rhizopus arrhizus* (B81a) were maintained at the CAMS Collection Center of Pathogen Microorganisms-D (CAMS-CCPM-D; Nanjing, China) and collected from MEA (Malt Extract Agar) media after a 4-day culture at 28°C. Genomic DNA, at least 1μg, was used for sequencing library construction. Paired-end libraries with insert sizes of ~400bp were prepared following Illumina's standard genomic DNA library preparation procedure. Purified genomic DNA was sheared into smaller fragments with a desired size by Covaris, and blunt ends were generated by using T4 DNA polymerase. The 3' end of the blunt phosphorylated DNA fragments was added with an 'A' base, then adapters were ligated to the ends of the DNA fragments. Through gel-electrophoresis, the desired fragments can be purified, selectively enriched, and amplified by PCR. The index tag could be introduced into the adapter at the PCR stage as appropriate. After a library quality test, the qualified Illumina pair-end library would be used for Illumina NovaSeq 6000 sequencing (150bp*2, Shanghai BIOZERON Co., Ltd.).

Strains genomic DNA was initially extracted using Invitrogen PureLink® Genomic DNA kit and tested by the NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific) for its quantity and quality. Purified further using the Quick-DNA Miniprep Plus kit, the whole genome sequencing was performed by Pacific Biosciences Sequel II technology (PacBio). Then the DNA SMRTbell libraries were made using the Express Template Prep Kit 2.0 from PacBio according to the manufacturer's protocol. Samples were pooled into a single multiplexed library and size was selected using Sage Sciences' BluePippin with the 0.75% DF Marker S1 High-Pass 6 - 10 kb v3 run protocol and S1 marker. Another size selection cut-off of 8000 (BPstart value) was used. After size selection, SMRTbell library was annealed and bound according to the SMRT Link Set Up and sequenced on a Sequel II. Raw PacBio reads were converted to fasta format with Samtools Fasta (http://www.htslib.org/doc/samtools.html).

## 2.3. Genome assembly and annotation

The raw paired end reads were trimmed, and quality-controlled by Trimmomatic with parameters (SLIDINGWINDOW: 4:15 MINLEN:75) (version 0.36 http://www.usadellab.org/cms/uploads/supplementary/Trimmomatic). Clean data obtained by above quality control processes was used to do further analysis.

The B50a, B63a, B66h and B81a genomes were sequenced using a combination of PacBio RS and Illumina sequencing platforms. The Illumina data was used to evaluate the complexity of the genome and correct the PacBio long reads. Firstly, we used ABySS (http://www.bcgsc.ca/platform/bioinfo/software/abyss) to do genome assembly with multiple-Kmer parameters and to get the optimal results of the assembly. Secondly canu (https://github.com/marbl/canu) was used to assemble the PacBio corrected long reads. Finally, GapCloser software was subsequently applied to fill up the remaining local inner gaps and correct the

single base polymorphism (https://sourceforge.net/projects/soapdenovo2/files/GapCloser/) for the final assembly results.

We used Repeatmasker (http://www.repeatmasker.org/) by default parameters to identify Genome repeat information. The ab initio prediction method was used to get gene models for strain B50a, B63a, B66h and B81a. Gene models were identified using Augustus (http://bioinf.uni-greifswald.de/augustus/binaries/). Then all gene models were blast against non-redundant (NR in NCBI) database, SwissProt (http://uniprot.org), KEGG (http://www.genome.jp/kegg/), and COG (http://www.ncbi.nlm.nih.gov/COG) to do functional annotation by blastp module. In addition, tRNAs were identified using the tRNAscan-SE (v1.23, http://lowelab.ucsc.edu/tRNAscan-SE) and rRNA were determined with the RNAmmer (v1.2, http://www.cbs.dtu.dk/services/RNAmmer/).

## 2.4. Pan-genome analysis and phylogenetic tree creation

Despite of the genomes of the four clinical isolates, we obtained genomic data of 43 more fungi from the National Center for Biotechnology Information's (NCBI) GenBank facility, including 25 species of *Mucorales* and other 22 pathogenic non-*Mucorales* species (Table S1). Core pan-genome analysis and evolutionary tree construction were carried out with the Orthofinder software (version 2.3.11) with default parameters. The information of *Mucorales*-specific and genus-specific orthogroups were extracted by Python script.

## 2.5. Virulence factors analysis

We have also selected proteins related to the virulence of *Mucorales*, referring to the published literature, and used blast v2.2.31 against the sequence of the 47 strains we included in total (Table 1). In this analysis, blast results were considered significant when the E-value was <1e-5, which is a generally accepted consensus cut-off. The best query results selected were combined with orthogroup gene assemblies to construct specific orthogroups related to virulence.

Table 1
Information of the genes related to the virulence of *Mucorales* from published literature.

| Name | Gene | GeneID | Platform | Description | Validated strains | Ref |
|---|---|---|---|---|---|---|
| *atf* | *atf1* | 190180 | JGP | activating transcription factor 1 | *Mucor circinelloides* | 6 |
| | *atf2* | 156289 | JGP | activating transcription factor 2 | Unknown | |
| *gcn4* | *gcn4* | 85517 | JGP | general control nonderepressible 4 | Unknown | |
| *aqp1* | *aqp1* | 167023 | JGP | aquaporin 1 | Unknown | |
| *ico1* | *ico1* | 155573 | JGP | isochorismatase 1 | Unknown | |
| *igp1* | *igp1* | 154866 | JGP | immunoglobulinlike protein | Unknown | |
| *pps1* | *pps1* | 115037 | JGP | signal peptide-containing protein | Unknown | |
| *arf* | *arf1* | 157293 | JGP | ADP-ribosylation factor | *Mucor circinelloides* | 1 |
| | *arl1* | 155647 | JGP | Arf-like proteins | Unknown | |
| | *arl2* | 112008 | JGP | Arf-like proteins | Unknown | |
| *mcmyo5* | *mcmyo5* | 51513 | JGP | Myosin class V protein | *Mucor circinelloides* | 5 |
| *mcplD* | *mcplD* | 134906 | JGP | Phospholipase D like protein | Unknown | |
| *112092* | *112092* | 112092 | JGP | hypothetical protein | *Mucor circinelloides* | 12 |
| *hmgR* | *hmgR1* | KJ508882 | NCBI | HMG-CoA reductase 1 gene, partial cds | *Mucor circinelloides f. lusitanicus*, *Aspergillus fumigatus*, *Candida albicans* or *Fusarium graminearum* | 2 |
| | *hmgR2* | KJ508884.1 | NCBI | HMG-CoA reductase 2 gene, partial cds | | |
| | *hmgR3* | KJ508883.1 | NCBI | HMG-CoA reductase 3 gene, partial cds | | |
| *cotH3* | *cotH3* | EIE87171.1 | NCBI | spore coat protein homologs RO3G_11882 | *Rhizopus delemar RA 99-880* | 9 |

JGP: Joint Genome Institute Mucor circinelloides v2.0 (http://genome.jgi.doe.gov/vista_embed/?organism = Mucci2)

| Name | Gene | GeneID | Platform | Description | Validated strains | Ref |
|---|---|---|---|---|---|---|
| *fob* | *fob2* | EIE86289.1 | NCBI | ferrioxamine binding cell surface protein RO3G_11000 | *Rhizopus delemar RA 99-880* | 11 |
| | *fob1* | EIE80382.1 | NCBI | ferrioxamine binding cell surface proteinsRO3G_05087 | *Rhizopus delemar RA 99-880* | |
| JGP: Joint Genome Institute Mucor circinelloides v2.0 (http://genome.jgi.doe.gov/vista_embed/?organism = Mucci2) | | | | | | |

NCBI: National Center for Biotechnology information (https://www.ncbi.nlm.nih.gov/protein)

# 3. Results

# 3.1. Causative pathogens of mucormycosis

The initial database searching results covered 672 articles with 25 duplications, which had been deleted with exclusions (Figure 1). 631 new patient datas from 272 eligible articles were included in the final analysis (Appendix S2). Merged with the data from 2000 to 2017 without duplications [4], 1078 individual cases all together have been analyzed.

A total of eight major genera (36 specific species and several unidentified species) of *Mucorales* organisms were identified in 1078 mucormycosis patients (Figure 2). Frequency outcomes showed that *Rhizopus* spp. (616/1078, 57.14%) was the most common pathogenic species, followed by *Mucor* spp. (124/1078, 11.5%) and *Lichtheimia* spp. (105/1078, 9.74%). So far, *Rhizopus arrhizus* (239/403, 59.31%) and *Lichtheimia corymbifera* (23/45, 51.11%) was the most frequent strain in *Rhizopus* spp. and *Lichtheimia* spp. respectively. Majority of *Mucor* spp. had not been identified to species level (48/61, 78.69%), which might be partially accounted for the countries with high incidence of mucormycosis lacking in adequate laboratory technologies for their poor sanitation and poverty. The top 10 identified pathogenic strains, in sequence, were *Rhizopus arrhizus* (28.66%), *Rhizopus microspores* (8.26%), *Lichtheimia corymbifera* (5.94%), *Cunninghamella bertholletiae* (3.53%), *Apophysomyces elegans* (3.06%), *Apophysomyces variabilis* (2.6%), *Rhizopus homothallicus* (2.41%), *Saksenaea vasiformis* (2.32%), *Rhizomucor pusillus* (1.76%), *Lichtheimia ramosa* (1.3%). To our surprise, *Mucor racemosus*, which had been recognized as one of the medically important *Mucorales* [20] showed no reported case in our data, as well as Jeong's result [4]. However, as Prakash has figured out, the true prevalence may be far underestimated in mucormycosis, with the difficulties in collecting specimens and diagnostic tests, especially in some high incidence Asian countries [8].

# 3.2. Sequence analysis and assemblies of the whole genomes

De novo sequencing of four pathogenic *Mucorales* isolates B50a, B63a, B66h and B81a (Table 2) were generated by Illumina Novaseq 6000 with PE150 mode. B81a had the longest raw reads as 4.88GB, while B66h had the shortest one as 13.7GB. B50a (8.11GB raw reads) and B63a (11.68GB raw reads) were intermediate in the sequence reads. Correspondingly, B50a had the minimum PacBio reads as 931,044,277 while B63a had the maximum ones as 1,546,237,919. In between, B66h had 1,167,129,942 and B81a had 1,403,589,167 PacBio reads. All four genomes had similar Q30 value of clean reads (92.73–93.97%). For following genome assembly and analysis, we have chosen the entire subreads longer than 500bp. B81a had the largest genome of the four isolates (51.80Mb), while B66h had the smallest one (32.29Mb). B50a (33.99Mb) and B63a (38.99Mb) owned the moderate genome sizes. The overall G+C content of the four isolates entire genome were 37.86% in B50a, 43.44% in B63a, 33.05% in B66h and 35.71% in B81a. We predicted 8,458 genes in B50a, 11,099 genes in B63a, 5,013 genes in B66h and 11,109 genes in B81a, which covered a total of 12,932,300 bp of the genome with a mean length of 1,529bp per gene, total of 16,636,830 bp of the genome with 1499bp per gene, total of 6,975,655bp of the genome with 1392 per gene, and total of 13,553,219bp of the genome with 1,220 per gene, respectively. Moreover, we identified 37 rRNA genes and 242 tRNA genes in B50a, 257 rRNA genes and 254 tRNA genes in B63a, 149 rRNA genes and 300 tRNA genes in B66h, and 40rRNA genes and 299 tRNA genes in B81a.

The whole genome sequencing of four *Mucorales* isolates have been uploaded to NCBI respectively: SRR16108680: *Mucor irregularis* (B50a), SRR16108679: *Lichtheimia corymbifera* (B63a), SRR16108678: *Mucor hiemalis* (B66h) and SRR16108677: *Rhizopus arrhizus* (B81a).

Table 2
De novo sequencing results of four pathogenic *Mucorales* isolates B50a, B63a, B66h and B81a.

| | *Mucor irregularis* | *Lichtheimia corymbifera* | *Mucor hiemalis* | *Rhizopus arrhizus* |
|---|---|---|---|---|
| Genome Size: | 33,988,018 | 38,994,095 | 32,290,736 | 51,800,396 |
| GC Content(%): | 37.87 | 43.45 | 33.05 | 35.7 |
| Gene Number: | 8,457 | 11,099 | 5,013 | 11,109 |
| Gene Total Length: | 12,930,096 | 16,636,830 | 6,975,655 | 13,553,220 |
| Gene Average Length: | 1,529 | 1,499 | 1,392 | 1,220 |
| Gene's GC Content: | 42.25 | 46.04 | 38.21 | 39.2 |
| % of Genome (Genes): | 38.04 | 42.66 | 21.6 | 26.16 |
| Intergenic region Length: | 21,057,922 | 22,357,265 | 25,315,081 | 38,247,176 |
| Intergenic's GC Content: | 35.18 | 41.52 | 31.63 | 34.46 |
| % of Genome (Intergenic): | 61.96 | 57.34 | 78.4 | 73.84 |

# 3.3. Repetitive sequences and genome annotation

Repetitive sequences (RS) have been shown to serve as vehicle to maintain genomic variability and to serve evolutionary change. We identified a total of 6,488 RSs in B50a occupying 2.12% of the global genome, 13,337 RSs in B63a with 3.44%, 6,214 RSs in B66h with 1.98%, and 20,845 RSs in B81a with 20.90%.

The gene sequences of B50a, B63a, B66h and B81a were compared against NR, GO, SWISS-Prot, KEGG and EGGNOG databases. The prediction of gene function from NR revealed 7692 genes which accounted for 90.95% of entire genome in B50a, 10938 genes in B60a with a percentage of 98.55%, 4800 genes in B66h with a percentage of 95.75% and 10689 genes in B81a took a percentage of 96.22%.

With NR protein database, the prediction of gene function from the clusters of orthogroups (KOG) database were 5541 genes in B50a, 6970 genes in B63a, 3847 genes in B66h and 7137 genes in B81a (data not shown).

# 3.4. Pan-genome and phylogenetic analysis

According to the pathogenic proportions of mucormycosis, we assembled, annotated and analyzed the genomes of (i) 25 species of the *Mucorales* order, including *Rhizopus* (*Rhizopus arrhizus*, *Rhizopus microsporus*, *Rhizopus stolonifer*, *Rhizopus delemar*), *Mucor* (*Mucor circinelloides*, *Mucor hiemalis*, *Mucor indicus*, *Mucor irregularis*, *Mucor lusitanicus*, *Mucor racemosus* and *Mucor velutinosus*), *Lichtheimia*

(*Lichtheimia corymbifera*, *Lichtheimia ramosa*), *Apophysomyces* (*Apophysomyces elegans*, *Apophysomyces ossiformis*, *Apophysomyces trapeziformis*, *Apophysomyces variabilis*), *Rhizomucor* (*Rhizomucor miehei*, *Rhizomucor pusillus*), *Cunninghamella* (*Cunninghamella bertholletiae*, *Cunninghamella elegans*), *Syncephalastrum* (*Syncephalastrum monosporum*, *Syncephalastrum racemosum*), *Actinomucor* (*Actinomucor elegans*) and *Cokeromyces* (*Cokeromyces recurvatus*), and (ii) 22 isolates of the non-*Mucorales* genera, including *Conidiobolus* (*Conidiobolus incongruous* and *Conidiobolus coronatus*), *Basidiobolus* (*Basidiobolus heterosporus* and *Basidiobolus meristosporus*), *Dermatophyte* (*Trichophyton rubrum*, *Trichophyton violaceum*, *Trichophyton mentagrophytes* and *Microsporum canis*), *Candida* (*Candida albicans*, *Candida glabrata*, *Candida parapsilosis*, *Candida tropicalis*, *Candida orthopsilosis* and *Candida metapsilosis*), *Cryptococcus* (*Cryptococcus neoformans* and *Cryptococcus gattii*), *Malassezia* (*Malassezia furfur*, *Malassezia dermatis* and *Malassezia globosa*), and *Aspergillus* (*Aspergillus niger*, *Aspergillus flavus* and *Aspergillus fumigatus*) to identify both of common and taxa-specific genetic elements in *Mucorales* and pathogenic non-*Mucorales* fungi (Table S2).

Pan-genome analysis was conducted to figure out *Mucorales*-specific genes, as well as in different *Mucorales* species to find out *Mucorales* genus-specific genes. One comparison of all 47 fungal genomes indicated that *Mucorales* species have 50 specific orthogroups which appear discrepant expression levels (Figure 3). *Mucorales*-specific orthogroups showed distinct homology to Non-Redundant (NR) Protein Sequence Database (Table 3). In sequence, OG0003655 had 96.9% similarity with EPB92269.1 as the STE/STE20 protein kinase, OG0002531 had 83% similarity with GAN03371.1 as TRM2 tRNA methyltransferase 2 homolog A, OG0000774 had 81.2% similarity with GAN00621.1 as pkinase-domain-containing protein, OG0004491 had 78.3% similarity with GAN01313.1 as transmembrane and coiled-coil domain-containing protein 3, and OG0002100 had 76.7% similarity with OAD09236.1 as glycoside hydrolase family 36 protein. Another comparison in the 9 *Mucorales* genera, genomes shared 1157 common orthogroups (Figure 4). *Cunninghamella* spp. owned the most specific orthogroups with 741 ones, followed by *Lichtheimia* spp. (333 orthogroups). Instead, *Rhizopus* spp. had 22 specific orthogroups and *Mucor* spp. had none. Based on the KOG database, the putative proteins were functionally classified into 22 molecular families covering cellular structure, biochemistry metabolism, molecular processing, and signal transduction (Figure 5). Orthogroups functioned more as signal transduction mechanisms in *Apophysomyces* spp. at 22.34% (42/188). The common orthogroups of *Cunninghamella* spp. gathered more in posttranslational modification, protein turnover, chaperones at 14.18% (37/261), as well as that of *Lichtheimia* spp. at 26.80% (26/97). Intracellular trafficking, secretion, and vesicular transport appeared to be main function in the common orthogroups of *Syncephalastrum* species.

Table 3
Information of the *Mucorales*-specific orthogroups compared to NR database.

| Orthogroup | NR tophit name | NR tophit description | NR top Similarity |
|---|---|---|---|
| OG0003655 | EPB92269.1 | STE/STE20 protein kinase [*Mucor circinelloides f. circinelloides* 1006PhL] | 96.9% |
| OG0002531 | GAN03371.1 | TRM2 tRNA methyltransferase 2 homolog A [*Mucor ambiguus*] | 83% |
| OG0000774 | GAN00621.1 | pkinase-domain-containing protein [*Mucor ambiguus*] | 81.2% |
| OG0004491 | GAN01313.1 | transmembrane and coiled-coil domain-containing protein 3 [*Mucor ambiguus*] | 78.3% |
| OG0002100 | OAD09236.1 | glycoside hydrolase family 36 protein [*Mucor circinelloides f. lusitanicus* CBS 277.49] | 76.7% |
| OG0001006 | GAN03654.1 | sel1 repeat protein [*Mucor ambiguus*] | 69.1% |
| OG0003562 | PHZ17618.1 | Quino protein alcohol dehydrogenase-like protein [*Rhizopus microsporus* ATCC 52813] | 67.5% |
| OG0002562 | GAN07947.1 | small GTPase rabE [*Mucor ambiguus*] | 58.6% |
| OG0002218 | OBZ90755.1 | Centrosomal protein [*Choanephora cucurbitarum*] | 51.1% |
| OG0002849 | OBZ88254.1 | DNA-dependent protein kinase catalytic subunit [*Choanephora cucurbitarum*] | 51% |

Based on the pan-genome analysis data, a maximum likelihood-based phylogenetic tree was constructed in 25 *Mucorales* fungi and 18 non-*Mucorales* fungi, with 4 *Entomopthorales* fungi as outgroups (Figure 6). We confirmed that *Mucorales* species were far away from non-*Mucorales* species. In addition, we noticed that the *M. irregularis* (B50a) and *M. hiemalis* (B66h) were clustered in one clade, sharing a very close relationship. *Cokeromyces recurvatus* and *Actinomucor elegans* were phylogenetically related to *Mucor* spp. within one phylogenetic cluster.

# 3.5. Expressions of proteins related to the virulence

To verify the reported proteins related to the virulence of *Mucorales* species, the copies of these proteins were compared in *Mucorales* species and other pathogenic fungi. Despite that *arf* and *aqp1* showed no difference, *atf*, *gcn4*, *igp1*, *pps1*, *mcmyo5*, *mcplD*, *112092*, *hmgR*, *cotH3* and *fob* had significant difference while *ico1* varied slightly between the two groups (Figure 7). For further knowledge of the unequal pathogenicity of the *Mucorales* species, we listed the expression of proteins related to the virulence in all the fungi we have referred above (Figure 8). Few expressions of *112092*, *cotH3*, *gcn4* and *igp1* was detected in non-*Mucorales* species, while almost all the *Mucorales* species expressed in different levels. Although the protein expression of *fob* and *hmgR* exhibited significant difference between *Mucorales* species and other pathogenic fungi, both were not as high as other proteins.

Compared to the non-*Mucorales* species, *Mucorales* species displayed abundant but uneven expression levels of *pps1*, *atf*, *mcplD* and *mcmyo5*. Gene *pps1* expressed mainly in *Rhizopus stolonifer*, *Rhizopus delemar* and *Mucor irregularis*; *atf* had high expression in *Mucor racemosus*, *Rhizopus delemar*, *Mucor lusitanicus*, *Mucor circinelloides* and *Rhizopus arrhizus*; *mcplD* expressed more in *Rhizopus stolonifer*, *Apophysomyces elegans* and *Mucor racemosus*; *mcmyo5* had multiple expressions in *Rhizopus stolonifer*, *Mucor racemosus* and *Rhizopus delemar.*

# 4. Discussion

Mucormycosis often causes high morbidity and mortality for its rapidly processive and destructive nature, of which urgent surgical and medical intervention can be lifesaving [21] [22] [23] [24]. Successful management depends on early diagnosis and optimal intervention. The availability of imaging techniques, trained and experienced physician, and mycological and histological investigations are extremely important in diagnosing mucormycosis. In the global guideline for the diagnosis and management of mucormycosis published in 2019, molecular-based methods for direct detection are moderately supported for the lack of standardization [25]. Several attempts have been done to detect *Mucorales* DNA in tissue samples [26–30], blood [31–33] and other fluids [34–37] as a non-invasive method of early diagnosis or pre-emptive therapy. PCR amplification of various DNA targets (ITS, 18S, 28S, cytochrome B) have been evaluated but lacking standardization for clinical application. Our study aimed to discover more potential DNA targets to enhance the possibility of molecular-based methods for rapid and accurate detection.

Based on the pan-genomic comparison of *Mucorales* species and non-*Mucorales* pathogenic fungi species, we have exhibited a few *Mucorales*-specific genes. EPB92269.1, encoding STE/STE20 protein kinase, seems to be a fundamental role in mitogen-activated protein kinase (MAPK) pathway, regulating cell wall maintenance, filamentous growth, and virulence [38]. Series of STE20 homologs have been suggested as a critical virulence factor for *Candida albicans* as CST20 [39] and CaCla4p [40], and *Trichophyton rubrum* as the p21 activated kinase (PAK) [41]. OAD09236.1 encodes glycoside hydrolase family 36 protein (GH36), which usually contains bacterial α-galactosidases. Up to now, a few fungal α-galactosidases that belong to GH36 have been reported, such as those from *Lichtheimia ramosa* (*LrAgal36A*) [42], *Rhizomucor miehei* (*RmGal36*) [43] and *Gibberella* sp. F75 (*aga-F75*) [44]. Different GH36 α-galactosidases show distinct activities under acidic conditions and physiological temperatures [44]. GAN03654.1 has high similarity with sel1 repeat protein, which has been suggested as the virulence determinant in *Legionella pneumophila* (*lpnE*), influencing vacuolar trafficking [45].

Seen from our frequency outcomes, the causative pathogens of mucormycosis with high incidence are *Rhizopus* species, *Mucor* species and *Lichtheimia* species. Although only sparse evidence supports that identification of the causative *Mucorales* to the genus or species level, or both, could guide the choice of the antifungal treatment [46–48], identification down to the species level is of importance for improved epidemiological knowledge in some outbreaks[49, 50]. However, identification of the *Mucorales* species in culture by standard mycological methods such as morphology is notoriously difficult because different

species share similar morphological characteristics (Figure 9). Some reference laboratories pointed out a high level of concordance (>90%) between morphology and molecular identification [51]. Also, some novel *Mucorales* PCR assays [52, 53] are released but likely to be designed to detect the order *Mucorales* as a whole, or individual genera, while *Mucorales* genus-specific gene targets still remain to be discussed.

In our pan-genome comparisons of *Mucorales* species, 22 specific orthogroups were detected but described only as hypothetical proteins for *Rhizopus* species which need to be further studied (data not shown). Beyond our expectation, none of the orthogroups was specific for *Mucor* species. Seven isolates had been included and compared in *Mucor* species, much more than other species, which might account for this unexpected outcome. When comparing the amino acid sequences specific in *Lichtheimia* species, several homologs show high similarity to identified proteins in posttranslational modification, protein turnover, chaperones, lipid transport and metabolism, including AMP-dependent synthetase ligase; fatty acid ligase; vacuolar transporter chaperone 4; a chain RMP-pepstatin A complex; DnaJ homolog subfamily A member 4; aspartic protease; cysteine proteinase; calmodulin; ubiquitin carboxyl-terminal hydrolase 12-like protein; multidrug resistance (MDR)-associated protein and MFS multidrug transporter (Table 4).

Table 4
*Lichtheimia* species-specific homologs.

| Orthogroup | NR tophit name | NR tophit description | NR top Similarity |
|---|---|---|---|
| OG0020222 | CDH52151.1 | achain rmp-pepstatin a complex [Lichtheimia corymbifera JMRC:FSU:9682] | 100 |
| OG0020221 | CDH52152.1 | achain x-ray analyses of aspartic structure andrefinement at angstroms resolution of the asparticproteinase from mucor pusillus [Lichtheimia corymbifera JMRC:FSU:9682] | 100 |
| OG0011126 | CDH58664.1 | achain x-ray analyses of aspartic structure andrefinement at angstroms resolution of the asparticproteinase from mucor pusillus [Lichtheimia corymbifera JMRC:FSU:9682] | 99.5 |
| OG0011105 | CDH50014.1 | amp binding protein [Lichtheimia corymbifera JMRC:FSU:9682] | 98.4 |
| OG0020099 | CDH50004.1 | amp-dependent synthetase ligase [Lichtheimia corymbifera JMRC:FSU:9682] | 100 |
| OG0020129 | CDH48850.1 | aspartic protease [Lichtheimia corymbifera JMRC:FSU:9682] | 100 |
| OG0020281 | CDH59824.1 | calmodulin [Lichtheimia corymbifera JMRC:FSU:9682] | 96 |
| OG0020241 | CDH56523.1 | cysteine proteinase [Lichtheimia corymbifera JMRC:FSU:9682] | 97.2 |
| OG0020248 | CDH58463.1 | dnaj homolog subfamily a member 4 [Lichtheimia corymbifera JMRC:FSU:9682] | 100 |
| OG0015056 | CDH49832.1 | fatty-acid–ligase [Lichtheimia corymbifera JMRC:FSU:9682] | 95.6 |
| OG0020268 | CDH49831.1 | long-chain-fatty-acid–ligase [Lichtheimia corymbifera JMRC:FSU:9682] | 100 |
| OG0020236 | CDH57209.1 | major facilitator superfamily protein [Lichtheimia corymbifera JMRC:FSU:9682] | 82.5 |
| OG0020075 | CDH57125.1 | mef2c protein [Lichtheimia corymbifera JMRC:FSU:9682] | 89.9 |
| OG0020042 | CDH50594.1 | mfs multidrug transporter [Lichtheimia corymbifera JMRC:FSU:9682] | 89.3 |

Additionally, *Cunninghamella* genera have been associated with an increased mortality rate in patients [46, 47] and shown to be more virulent in experimental models[48], with several homologs shown high similarity in Posttranslational modification, protein turnover, chaperones and transcription. Such as vanillate O-demethylase; secretory lipase, glutathione S-transferase; alpha/beta hydrolase fold family; ferric reductase, NADH/NADPH oxidase; amino acid permease/ SLC12A domain-containing protein; thiolase; C6 transcription factor; DCG1-like protein; exo-beta-1; pantothenate transporter; cathepsin L-like; FAD binding domain protein and so on (Table 5).

| Orthogroup | NR tophit name | NR tophit description | NR top Similarity |
|---|---|---|---|
| OG0020122 | CDH49091.1 | multidrug resistance-associated protein 1-like [Lichtheimia corymbifera JMRC:FSU:9682] | 98.3 |
| OG0020150 | CDH60432.1 | ubiquitin carboxyl-terminal hydrolase 12-like [Lichtheimia corymbifera JMRC:FSU:9682] | 93.4 |
| OG0020186 | CDH60762.1 | vtc4p [Lichtheimia corymbifera JMRC:FSU:9682] | 100 |

Additionally, *Cunninghamella* genera have been associated with an increased mortality rate in patients [46, 47] and shown to be more virulent in experimental models[48], with several homologs shown high similarity in Posttranslational modification, protein turnover, chaperones and transcription. Such as vanillate O-demethylase; secretory lipase, glutathione S-transferase; alpha/beta hydrolase fold family; ferric reductase, NADH/NADPH oxidase; amino acid permease/ SLC12A domain-containing protein; thiolase; C6 transcription factor; DCG1-like protein; exo-beta-1; pantothenate transporter; cathepsin L-like; FAD binding domain protein and so on (Table 5).

Table 5
*Cunninghamella* species-specific homologs.

| Orthogroup | NR tophit name | NR tophit description | NR top Similarity |
|---|---|---|---|
| OG0013781 | RMZ47002.1 | alpha/beta fold family hydrolase [Aspergillus flavus] | 100 |
| OG0001523 | KAB8248670.1 | amino acid permease/ SLC12A domain-containing protein [Aspergillus flavus] | 99.6 |
| OG0019965 | OAQ35848.1 | arsenate reductase [Mortierella elongata AG-77] | 62.8 |
| OG0010164 | RMZ48341.1 | Asp-hemolysin precursor [Aspergillus flavus] | 100 |
| OG0004525 | EED49147.1 | C6 transcription factor (Leu3), putative [Aspergillus flavus NRRL3357] | 86.1 |
| OG0005683 | KOC11690.1 | C6 transcription factor [Aspergillus flavus AF70] | 92.8 |
| OG0019923 | OQR66474.1 | cathepsin L-like [Tropilaelaps mercedesae] | 69 |
| OG0006743 | CAF32059.1 | DCG1-like protein, putative [Aspergillus fumigatus] | 99 |
| OG0019762 | PIA12612.1 | DNA polymerase [Coemansia reversa NRRL 1564] | 68.9 |
| OG0005197 | RMZ39367.1 | exo-beta-1 [Aspergillus flavus] | 98 |
| OG0007683 | PIG88081.1 | FAD binding domain protein [Aspergillus arachidicola] | 67.4 |
| OG0005204 | ORX90396.1 | FAD-binding domain-containing protein [Basidiobolus meristosporus CBS 931.73] | 74.2 |
| OG0000278 | KDE76689.1 | ferric reductase, NADH/NADPH oxidase [Aspergillus oryzae 100-8] | 99.9 |
| OG0000667 | EED56953.1 | glutathione S-transferase, putative [Aspergillus flavus NRRL3357] | 100 |
| OG0007862 | EAL86257.1 | GNAT family N-acetyltransferase, putative [Aspergillus fumigatus Af293] | 88.4 |
| OG0008586 | RAQ76907.1 | O-methyltransferase [Aspergillus flavus] | 100 |
| OG0006933 | OOO10681.1 | Oxidoreductase FAD-binding domain protein [Aspergillus oryzae] | 92.3 |
| OG0006958 | RMZ44899.1 | pantothenate transporter [Aspergillus flavus] | 84.5 |

According to the reported proteins related to virulence of *Mucorales* species, *112092*, *cotH3*, *gcn4* and *igp1* expressed few in non-*Mucorales* species but abundant in almost all *Mucorales* species. Studied from the former studys, deletion of gene *112092* results in significant reduction of *Mucor circinelloides* virulence [13]; *cotH3* has been identified as the fungal ligands that mediate attachment to *GRP78* during host cell invasion [18]; *gcn4* deletion causes severe germination and growth defects in the spores after phagocytosis, affecting the development of mucormycosis in mice [11]; deletion mutants in *igp1* develops healthy colonies similar to wild-type virulent strain but leads to less death of mice [11]. Thus, the four DNA targets might be designed for the direct identification as well as the virulence detection of *Mucorales* species.

| Orthogroup | NR tophit name | NR tophit description | NR top Similarity |
|---|---|---|---|
| OG0019721 | ORZ05233.1 | phosphoesterase family-domain-containing protein [Absidia repens] | 74.3 |
| OG0001102 | RMZ37442.1 | secretory lipase [Aspergillus flavus] | 100 |
| OG0006483 | KJK61788.1 | specific transcription factor domain protein [Aspergillus parasiticus SU-1] | 93.5 |
| OG0006748 | RDH19797.1 | squalene-hopene-cyclase [Aspergillus niger ATCC 13496] | 100 |
| OG0019940 | ORZ18196.1 | Thiolase, N-terminal domain-domain-containing protein [Absidia repens] | 93.8 |
| OG0009144 | RMZ46722.1 | vanillate O-demethylase oxidoreductase [Aspergillus flavus] | 100 |

According to the reported proteins related to virulence of *Mucorales* species, *112092*, *cotH3*, *gcn4* and *igp1* expressed few in non-*Mucorales* species but abundant in almost all *Mucorales* species. Studied from the former studys, deletion of gene *112092* results in significant reduction of *Mucor circinelloides* virulence [13]; *cotH3* has been identified as the fungal ligands that mediate attachment to *GRP78* during host cell invasion [18]; *gcn4* deletion causes severe germination and growth defects in the spores after phagocytosis, affecting the development of mucormycosis in mice [11]; deletion mutants in *igp1* develops healthy colonies similar to wild-type virulent strain but leads to less death of mice [11]. Thus, the four DNA targets might be designed for the direct identification as well as the virulence detection of *Mucorales* species.

# 5. Conclusions

We completed the whole genome sequencing of three common pathogenic *Mucorales* species and one uncommon species from CAMS-CCPM-D to study our epidemiological results of mucormycosis. As we have mentioned above, some regions cannot identify the *Mucorales* to genus or species which may lead to some deviation in our epidemiological statistics.

To find out the *Mucorales*-specific genes or genus-specific genes in *Mucorales*, and to seek the possibility of direct detections in serum or other body fluids for identification to genus/species, we compared the whole genome sequencing of *Mucorales* species and non-*Mucorales* species. A few *Mucorales*-specific genes have been found which are related to STE/STE20 protein kinase, GH36 and sel1 repeat protein. *Mucorales* genus-specific genes are also discussed here, especially those of *Lichtheimia* species and *Cunninghamella* species, which covered cellular structure, biochemistry metabolism, molecular processing, and signal transduction. Reported proteins related to the virulence of *Mucorales* species were varied with distinct significance while *hmgR3, 112092*, *cotH3*, *gcn4* and *igp1* have the potential to be the detection targets.

The best informative DNA target should have a large intergenus (between genus) and a low intragenus (within a given genus) sequence variability, which brings us the dawn of new technology for rapid and

convenient diagnosis, and accurate and efficient treatments.

# Declarations

**Author Contributions:** Conceptualization, Zhang, M.J. and Liang, G.Z.; methodology, Zhang, M.J. Xu, W.Q. and Liang, G.Z.; software, Zhang, M.J. and Tao, Y.; validation, Zhang, M.J. and Ge, N.C.; formal analysis, Zhang, M.J. and Song, G.; investigation, Zhang, M.J. and Liang, G.Z.; resources, Zhang, M.J. and Mei H.; data curation, Zhang, M.J.; writing—original draft preparation, Zhang, M.J.; writing—review and editing, Zhang, M.J. and Liang, G.Z.; visualization, Zhang, M.J.; supervision, Liu W.D. and Liang; project administration, Liu W.D. and Liang, G.Z.; funding acquisition, Liu W.D. and Liang, G.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the National Natural Science Foundation of China, grant number 81471905; the Nanjing Incubation Program for National Clinical Research Center, grant number 2019060001; the National Science and Technology Infrastructure of China, grant number National Pathogen Resource Center- NPRC- 32.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The whole genome sequencing of four *Mucorales* isolates have been uploaded to NCBI (https://www.ncbi.nlm.nih.gov) respectively: SRR16108680: *Mucor irregularis* (B50a), SRR16108679: *Lichtheimia corymbifera* (B63a), SRR16108678: *Mucor hiemalis* (B66h) and SRR16108677: *Rhizopus arrhizus* (B81a).

**Acknowledgments:** We would like to thank Shanghai BIOZERON Biotechnology Co. (www.biozeron.com) for providing sequencing technologies and bioinformatic analysis. Thanks for the wonderful supporting work from Fuxing Zha, Xiaomeng Cao and Yan Cheng.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

# References

1. Park BJ, Pappas PG, Wannemuehler KA, Alexander BD, Anaissie EJ, Andes DR, Baddley JW, Brown JM, Brumble LM, Freifeld AG *et al*: **Invasive non-Aspergillus mold infections in transplant recipients, United States, 2001-2006**. *Emerg Infect Dis* 2011, **17**(10):1855-1864.
2. Chakrabarti A, Chatterjee SS, Das A, Panda N, Shivaprakash MR, Kaur A, Varma SC, Singhi S, Bhansali A, Sakhuja V: **Invasive zygomycosis in India: experience in a tertiary care hospital**. *Postgrad Med J* 2009, **85**(1009):573-581.

3. Lelievre L, Garcia-Hermoso D, Abdoul H, Hivelin M, Chouaki T, Toubas D, Mamez AC, Lantieri L, Lortholary O, Lanternier F: **Posttraumatic mucormycosis: a nationwide study in France and review of the literature**. *Medicine (Baltimore)* 2014, **93**(24):395-404.
4. Jeong W, Keighley C, Wolfe R, Lee WL, Slavin MA, Kong DCM, Chen SC: **The epidemiology and clinical manifestations of mucormycosis: a systematic review and meta-analysis of case reports**. *Clin Microbiol Infect* 2019, **25**(1):26-34.
5. Prakash H, Ghosh AK, Rudramurthy SM, Singh P, Xess I, Savio J, Pamidimukkala U, Jillwin J, Varma S, Das A *et al*: **A prospective multicenter study on mucormycosis in India: Epidemiology, diagnosis, and treatment**. *Med Mycol* 2019, **57**(4):395-402.
6. Chakrabarti A, Das A, Mandal J, Shivaprakash MR, George VK, Tarai B, Rao P, Panda N, Verma SC, Sakhuja V: **The rising trend of invasive zygomycosis in patients with uncontrolled diabetes mellitus**. *Med Mycol* 2006, **44**(4):335-342.
7. Lin E, Moua T, Limper AH: **Pulmonary mucormycosis: clinical features and outcomes**. *Infection* 2017, **45**(4):443-448.
8. Prakash H, Chakrabarti A: **Global Epidemiology of Mucormycosis**. *J Fungi (Basel)* 2019, **5**(1).
9. Singh AK, Singh R, Joshi SR, Misra A: **Mucormycosis in COVID-19: A systematic review of cases reported worldwide and in India**. *Diabetes Metab Syndr* 2021, **15**(4):102146.
10. Song G, Liang G, Liu W: **Fungal Co-infections Associated with Global COVID-19 Pandemic: A Clinical and Diagnostic Perspective from China**. *Mycopathologia* 2020, **185**(4):599-606.
11. Pérez-Arques C, Navarro-Mendoza MI, Murcia L, Lax C, Martínez-García P, Heitman J, Nicolás FE, Garre V: **Mucor circinelloides Thrives inside the Phagosome through an Atf-Mediated Germination Pathway**. *mBio* 2019, **10**(1).
12. Trieu TA, Navarro-Mendoza MI, Pérez-Arques C, Sanchis M, Capilla J, Navarro-Rodriguez P, Lopez-Fernandez L, Torres-Martínez S, Garre V, Ruiz-Vázquez RM *et al*: **RNAi-Based Functional Genomics Identifies New Virulence Determinants in Mucormycosis**. *PLoS Pathog* 2017, **13**(1):e1006150.
13. López-Fernández L, Sanchis M, Navarro-Rodríguez P, Nicolás FE, Silva-Franco F, Guarro J, Garre V, Navarro-Mendoza MI, Pérez-Arques C, Capilla J: **Understanding Mucor circinelloides pathogenesis by comparative genomics and phenotypical studies**. *Virulence* 2018, **9**(1):707-720.
14. Patiño-Medina JA, Maldonado-Herrera G, Pérez-Arques C, Alejandre-Castañeda V, Reyes-Mares NY, Valle-Maldonado MI, Campos-García J, Ortiz-Alvarado R, Jácome-Galarza IE, Ramírez-Díaz MI *et al*: **Control of morphology and virulence by ADP-ribosylation factors (Arf) in Mucor circinelloides**. *Curr Genet* 2018, **64**(4):853-869.
15. Patiño-Medina JA, Valle-Maldonado MI, Maldonado-Herrera G, Pérez-Arques C, Jácome-Galarza IE, Díaz-Pérez C, Díaz-Pérez AL, Araiza-Cervantes CA, Villagomez-Castro JC, Campos-García J *et al*: **Role of Arf-like proteins (Arl1 and Arl2) of Mucor circinelloides in virulence and antifungal susceptibility**. *Fungal Genet Biol* 2019, **129**:40-51.
16. Nagy G, Vaz AG, Szebenyi C, Takó M, Tóth EJ, Csernetics Á, Bencsik O, Szekeres A, Homa M, Ayaydin F *et al*: **CRISPR-Cas9-mediated disruption of the HMG-CoA reductase genes of Mucor circinelloides**

**and subcellular localization of the encoded enzymes**. *Fungal Genet Biol* 2019, **129**:30-39.

17. Bellanger AP, Tatara AM, Shirazi F, Gebremariam T, Albert ND, Lewis RE, Ibrahim AS, Kontoyiannis DP: **Statin Concentrations Below the Minimum Inhibitory Concentration Attenuate the Virulence of Rhizopus oryzae**. *J Infect Dis* 2016, **214**(1):114-121.
18. Gebremariam T, Liu M, Luo G, Bruno V, Phan QT, Waring AJ, Edwards JE, Jr., Filler SG, Yeaman MR, Ibrahim AS: **CotH3 mediates fungal invasion of host cells during mucormycosis**. *J Clin Invest* 2014, **124**(1):237-250.
19. Liu M, Lin L, Gebremariam T, Luo G, Skory CD, French SW, Chou TF, Edwards JE, Jr., Ibrahim AS: **Fob1 and Fob2 Proteins Are Virulence Determinants of Rhizopus oryzae via Facilitating Iron Uptake from Ferrioxamine**. *PLoS Pathog* 2015, **11**(5):e1004842.
20. Singh P, Paul S, Shivaprakash MR, Chakrabarti A, Ghosh AK: **Stress response in medically important Mucorales**. *Mycoses* 2016, **59**(10):628-635.
21. Chamilos G, Lewis RE, Kontoyiannis DP: **Delaying amphotericin B-based frontline therapy significantly increases mortality among patients with hematologic malignancy who have zygomycosis**. *Clin Infect Dis* 2008, **47**(4):503-509.
22. Vaughan C, Bartolo A, Vallabh N, Leong SC: **A meta-analysis of survival factors in rhino-orbital-cerebral mucormycosis-has anything changed in the past 20 years?***Clin Otolaryngol* 2018, **43**(6):1454-1464.
23. Sun HY, Singh N: **Mucormycosis: its contemporary face and management strategies**. *Lancet Infect Dis* 2011, **11**(4):301-311.
24. Liang GZ, Xu WQ, Zheng XL, Mei H, Lv GX, Shen YN, Li DM, Liu WD: **Successful Treatment by Surgery of a Primary Cutaneous Mucormycosis Caused by Mucor irregularis**. *Mycopathologia* 2018, **183**(2):445-449.
25. Cornely OA, Alastruey-Izquierdo A, Arenz D, Chen SCA, Dannaoui E, Hochhegger B, Hoenigl M, Jensen HE, Lagrou K, Lewis RE *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**. *Lancet Infect Dis* 2019, **19**(12):e405-e421.
26. Guinea J, Escribano P, Vena A, Muñoz P, Martínez-Jiménez MDC, Padilla B, Bouza E: **Increasing incidence of mucormycosis in a large Spanish hospital from 2007 to 2015: Epidemiology and microbiological characterization of the isolates**. *PLoS One* 2017, **12**(6):e0179136.
27. Rickerts V, Mousset S, Lambrecht E, Tintelnot K, Schwerdtfeger R, Presterl E, Jacobi V, Just-Nübling G, Bialek R: **Comparison of histopathological analysis, culture, and polymerase chain reaction assays to detect invasive mold infections from biopsy specimens**. *Clin Infect Dis* 2007, **44**(8):1078-1083.
28. Alanio A, Garcia-Hermoso D, Mercier-Delarue S, Lanternier F, Gits-Muselli M, Menotti J, Denis B, Bergeron A, Legrand M, Lortholary O *et al*: **Molecular identification of Mucorales in human tissues: contribution of PCR electrospray-ionization mass spectrometry**. *Clin Microbiol Infect* 2015, **21**(6):594.e591-595.

29. Dannaoui E, Schwarz P, Slany M, Loeffler J, Jorde AT, Cuenca-Estrella M, Hauser PM, Shrief R, Huerre M, Freiberger T *et al*: **Molecular detection and identification of zygomycetes species from paraffin-embedded tissues in a murine model of disseminated zygomycosis: a collaborative European Society of Clinical Microbiology and Infectious Diseases (ESCMID) Fungal Infection Study Group (EFISG) evaluation**. *J Clin Microbiol* 2010, **48**(6):2043-2046.
30. Gholinejad-Ghadi N, Shokohi T, Seifi Z, Aghili SR, Roilides E, Nikkhah M, Pormosa R, Karami H, Larjani LV, Ghasemi M *et al*: **Identification of Mucorales in patients with proven invasive mucormycosis by polymerase chain reaction in tissue samples**. *Mycoses* 2018, **61**(12):909-915.
31. Kasai M, Harrington SM, Francesconi A, Petraitis V, Petraitiene R, Beveridge MG, Knudsen T, Milanovich J, Cotton MP, Hughes J *et al*: **Detection of a molecular biomarker for zygomycetes by quantitative PCR assays of plasma, bronchoalveolar lavage, and lung tissue in a rabbit model of experimental pulmonary zygomycosis**. *J Clin Microbiol* 2008, **46**(11):3690-3702.
32. Legrand M, Gits-Muselli M, Boutin L, Garcia-Hermoso D, Maurel V, Soussi S, Benyamina M, Ferry A, Chaussard M, Hamane S *et al*: **Detection of Circulating Mucorales DNA in Critically Ill Burn Patients: Preliminary Report of a Screening Strategy for Early Diagnosis and Treatment**. *Clin Infect Dis* 2016, **63**(10):1312-1317.
33. Kobayashi M, Togitani K, Machida H, Uemura Y, Ohtsuki Y, Taguchi H: **Molecular polymerase chain reaction diagnosis of pulmonary mucormycosis caused by Cunninghamella bertholletiae**. *Respirology* 2004, **9**(3):397-401.
34. Shigemura T, Nakazawa Y, Matsuda K, Motobayashi M, Saito S, Koike K: **Evaluation of Mucorales DNA load in cerebrospinal fluid in a patient with possible cerebral mucormycosis treated with intravenous liposomal amphotericin B**. *Int J Infect Dis* 2014, **29**:200-202.
35. Lengerova M, Racil Z, Hrncirova K, Kocmanova I, Volfova P, Ricna D, Bejdak P, Moulis M, Pavlovsky Z, Weinbergerova B *et al*: **Rapid detection and identification of mucormycetes in bronchoalveolar lavage samples from immunocompromised patients with pulmonary infiltrates by use of high-resolution melt analysis**. *J Clin Microbiol* 2014, **52**(8):2824-2828.
36. Scherer E, Iriart X, Bellanger AP, Dupont D, Guitard J, Gabriel F, Cassaing S, Charpentier E, Guenounou S, Cornet M *et al*: **Quantitative PCR (qPCR) Detection of Mucorales DNA in Bronchoalveolar Lavage Fluid To Diagnose Pulmonary Mucormycosis**. *J Clin Microbiol* 2018, **56**(8).
37. Springer J, White PL, Kessel J, Wieters I, Teschner D, Korczynski D, Liebregts T, Cornely OA, Schwartz S, Elgeti T *et al*: **A Comparison of Aspergillus and Mucorales PCR Testing of Different Bronchoalveolar Lavage Fluid Fractions from Patients with Suspected Invasive Pulmonary Fungal Disease**. *J Clin Microbiol* 2018, **56**(2).
38. Boyce KJ, Andrianopoulos A: **Ste20-related kinases: effectors of signaling and morphogenesis in fungi**. *Trends Microbiol* 2011, **19**(8):400-410.
39. Leberer E, Harcus D, Broadbent ID, Clark KL, Dignard D, Ziegelbauer K, Schmidt A, Gow NA, Brown AJ, Thomas DY: **Signal transduction through homologs of the Ste20p and Ste7p protein kinases can**

**trigger hyphal formation in the pathogenic fungus Candida albicans**. *Proc Natl Acad Sci U S A* 1996, **93**(23):13217-13222.

40. Leberer E, Ziegelbauer K, Schmidt A, Harcus D, Dignard D, Ash J, Johnson L, Thomas DY: **Virulence and hyphal formation of Candida albicans require the Ste20p-like protein kinase CaCla4p**. *Curr Biol* 1997, **7**(8):539-546.
41. Gomes EV, Bortolossi JC, Sanches PR, Mendes NS, Martinez-Rossi NM, Rossi A: **STE20/PAKA Protein Kinase Gene Releases an Autoinhibitory Domain through Pre-mRNA Alternative Splicing in the Dermatophyte Trichophyton rubrum**. *Int J Mol Sci* 2018, **19**(11).
42. Xie J, Wang B, He Z, Pan L: **A thermophilic fungal GH36 α-galactosidase from Lichtheimia ramosa and its synergistic hydrolysis of locust bean gum**. *Carbohydr Res* 2020, **491**:107911.
43. Katrolia P, Jia H, Yan Q, Song S, Jiang Z, Xu H: **Characterization of a protease-resistant α-galactosidase from the thermophilic fungus Rhizomucor miehei and its application in removal of raffinose family oligosaccharides**. *Bioresour Technol* 2012, **110**:578-586.
44. Cao Y, Wang Y, Meng K, Bai Y, Shi P, Luo H, Yang P, Zhou Z, Zhang Z, Yao B: **A novel protease-resistant alpha-galactosidase with high hydrolytic activity from Gibberella sp. F75: gene cloning, expression, and enzymatic characterization**. *Appl Microbiol Biotechnol* 2009, **83**(5):875-884.
45. Newton HJ, Sansom FM, Dao J, McAlister AD, Sloan J, Cianciotto NP, Hartland EL: **Sel1 repeat protein LpnE is a Legionella pneumophila virulence determinant that influences vacuolar trafficking**. *Infect Immun* 2007, **75**(12):5575-5585.
46. Roden MM, Zaoutis TE, Buchanan WL, Knudsen TA, Sarkisova TA, Schaufele RL, Sein M, Sein T, Chiou CC, Chu JH *et al*: **Epidemiology and outcome of zygomycosis: a review of 929 reported cases**. *Clin Infect Dis* 2005, **41**(5):634-653.
47. Gomes MZ, Lewis RE, Kontoyiannis DP: **Mucormycosis caused by unusual mucormycetes, non-Rhizopus, -Mucor, and -Lichtheimia species**. *Clin Microbiol Rev* 2011, **24**(2):411-445.
48. Petraitis V, Petraitiene R, Antachopoulos C, Hughes JE, Cotton MP, Kasai M, Harrington S, Gamaletsou MN, Bacher JD, Kontoyiannis DP *et al*: **Increased virulence of Cunninghamella bertholletiae in experimental pulmonary mucormycosis: correlation with circulating molecular biomarkers, sporangiospore germination and hyphal metabolism**. *Med Mycol* 2013, **51**(1):72-82.
49. Etienne KA, Gillece J, Hilsabeck R, Schupp JM, Colman R, Lockhart SR, Gade L, Thompson EH, Sutton DA, Neblett-Fanfair R *et al*: **Whole genome sequence typing to investigate the Apophysomyces outbreak following a tornado in Joplin, Missouri, 2011**. *PLoS One* 2012, **7**(11):e49989.
50. Garcia-Hermoso D, Criscuolo A, Lee SC, Legrand M, Chaouat M, Denis B, Lafaurie M, Rouveau M, Soler C, Schaal JV *et al*: **Outbreak of Invasive Wound Mucormycosis in a Burn Unit Due to Multiple Strains of Mucor circinelloides f. circinelloides Resolved by Whole-Genome Sequencing**. *mBio* 2018, **9**(2).
51. Alvarez E, Sutton DA, Cano J, Fothergill AW, Stchigel A, Rinaldi MG, Guarro J: **Spectrum of zygomycete species identified in clinically significant specimens in the United States**. *J Clin Microbiol* 2009, **47**(6):1650-1656.

52. Guegan H, Iriart X, Bougnoux ME, Berry A, Robert-Gangneux F, Gangneux JP: **Evaluation of MucorGenius® mucorales PCR assay for the diagnosis of pulmonary mucormycosis**. *J Infect* 2020, **81**(2):311-317.
53. Rocchi S, Scherer E, Mengoli C, Alanio A, Botterel F, Bougnoux ME, Bretagne S, Cogliati M, Cornu M, Dalle F *et al*: **Interlaboratory evaluation of Mucorales PCR assays for testing serum specimens: A study by the fungal PCR Initiative and the Modimucor study group**. *Med Mycol* 2021, **59**(2):126-138.

# Figures

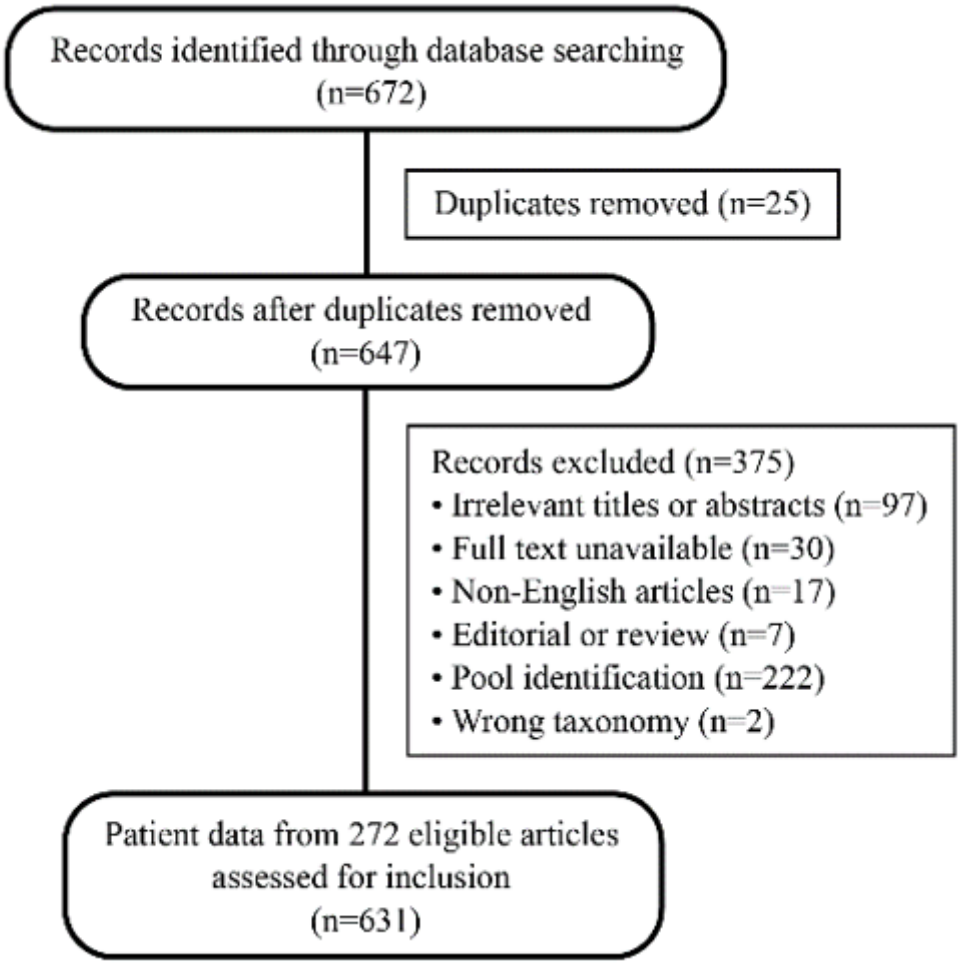
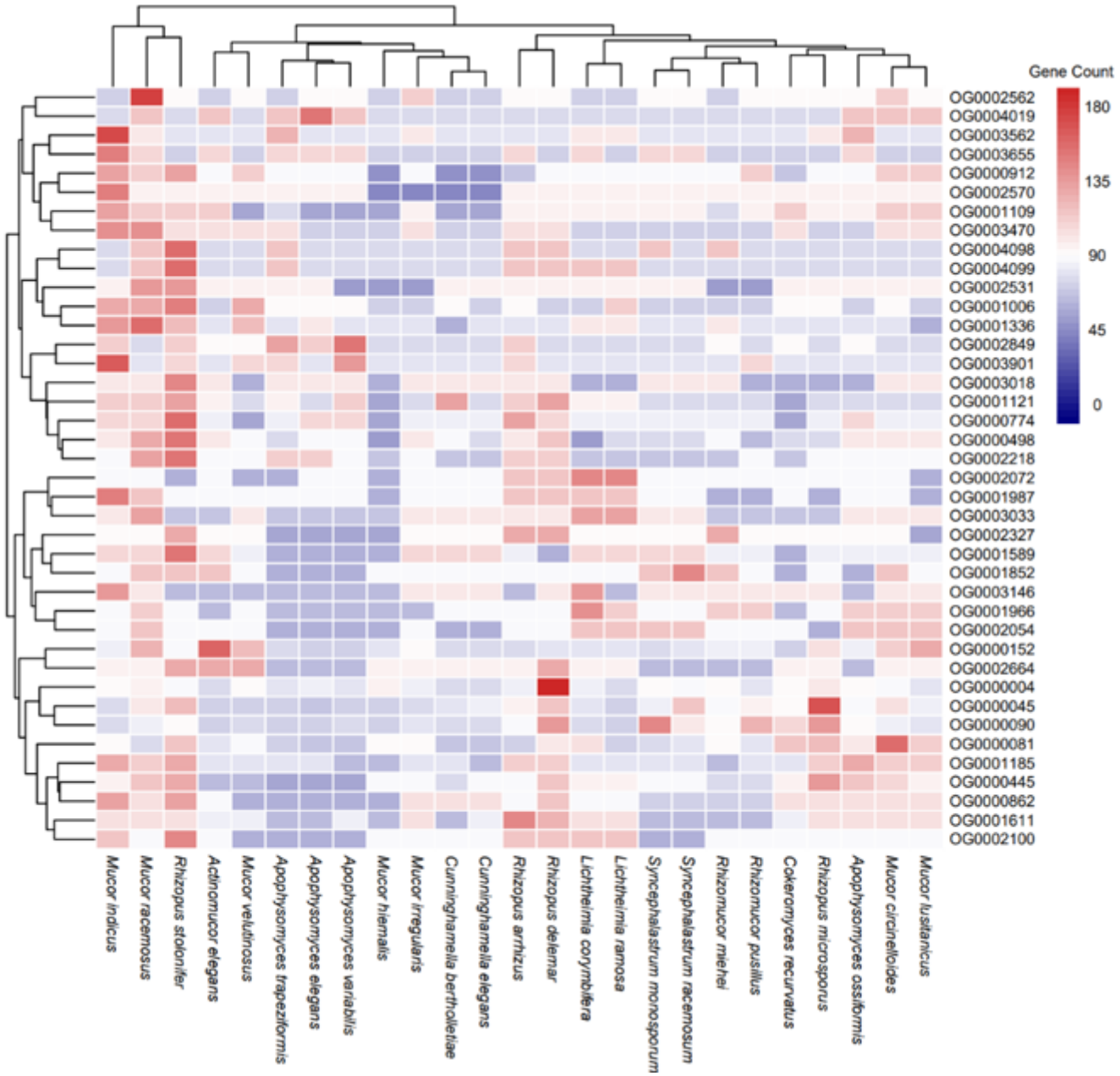



**Figure 1**

Searching flowchart for causative pathogens. There were 672 records received initially, of which removed the duplications (n=25) and records with exclusions (n=375). All in all, 631 new patient data from 272 eligible articles were included in our analysis.

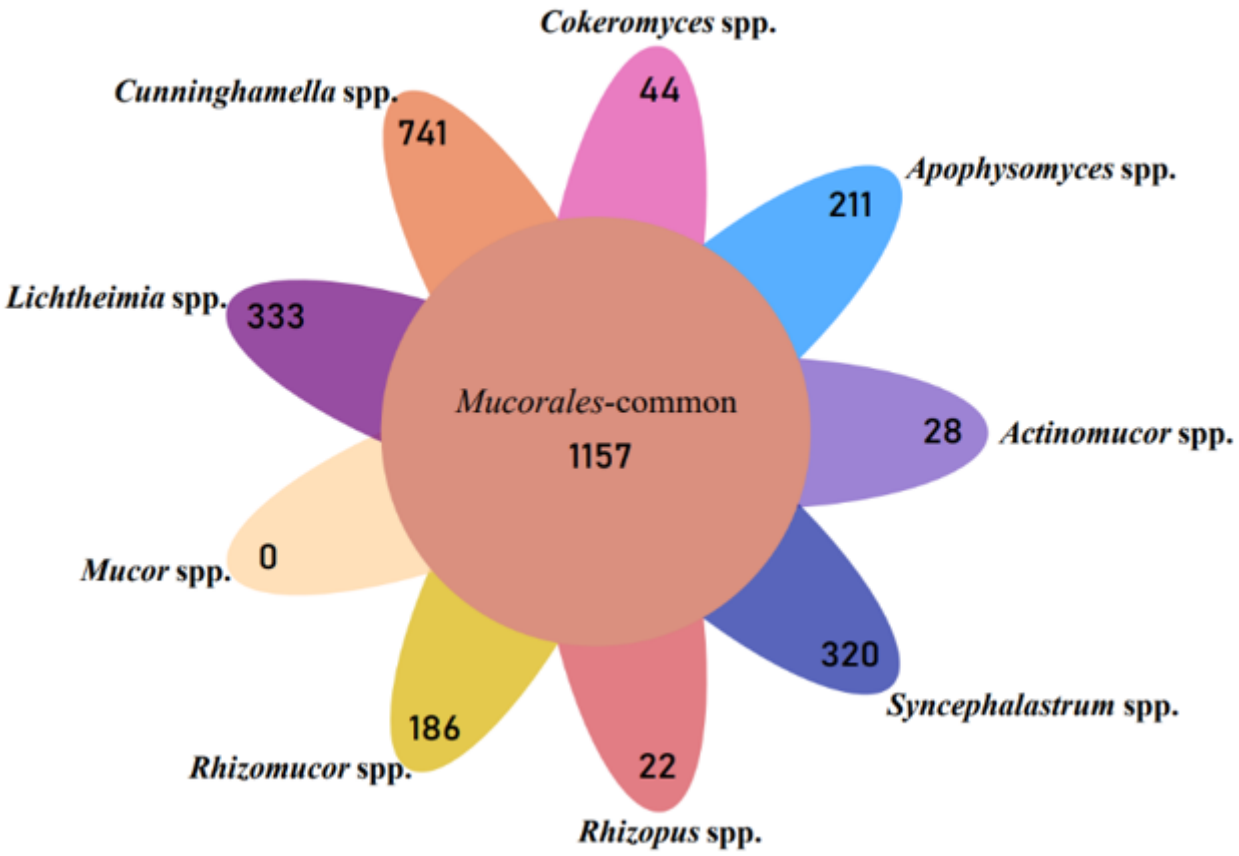
**Figure 2**

Causative pathogens condition of murcomycosis: (A) Proportions of different Mucorales genus, and (B) Proportions of different Mucorales species shown in the column diagram while the top 10 species individually pictured in pie charts.



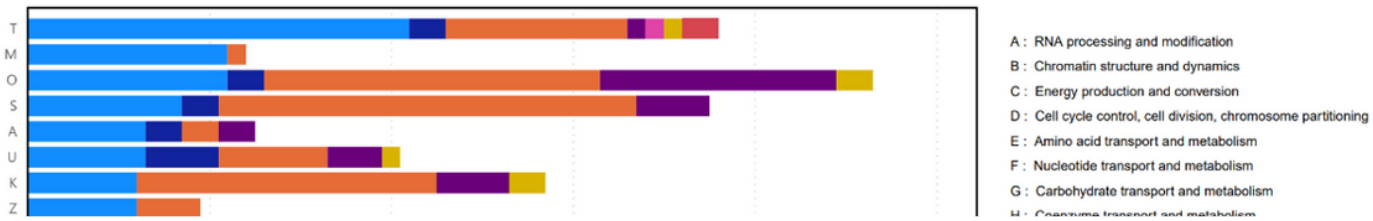

## Figure 3

There were 50 *Mucorales* genus-specific orthogroups appearing discrepant expression levels in different *Mucorales* species.




## Figure 4

Pan-genome comparison in the 9 *Mucorales* genera. All genomes shared 1157 common orthogroups and distinct specific orthogroups for each *Mucorales* genus except *Mucor* species.

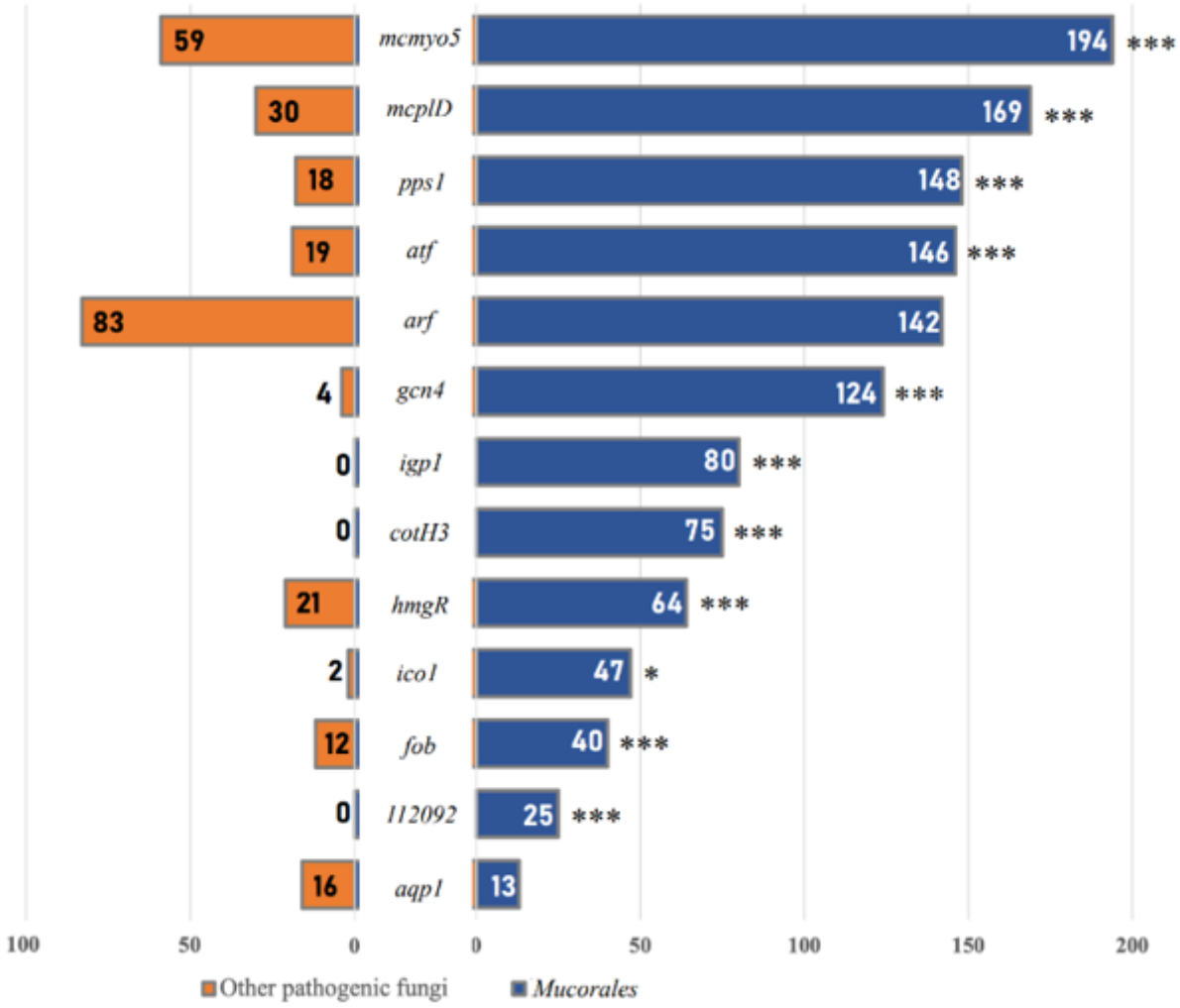


## Figure 5

Histogram of KOG classification of the 9 *Mucorales* genera. All unigenes were aligned to KOG database for prediction and classification based on available and possible functions.

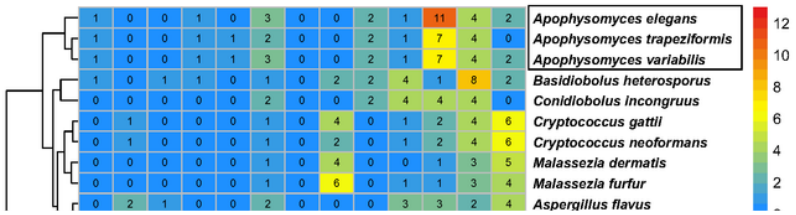
## Figure 6

Core pan analysis and evolutionary tree. Genera of *Mucorales* have been marked in different colors.




## Figure 7

Comparison of virulence related proteins between *Mucorales* species and other pathogenic fungi. (*: <0.05; ***: <0.001)




## Figure 8

Expression heatmap of proteins related to the virulence in all the 47 fungi. *Mucorales* species were in the black frames.

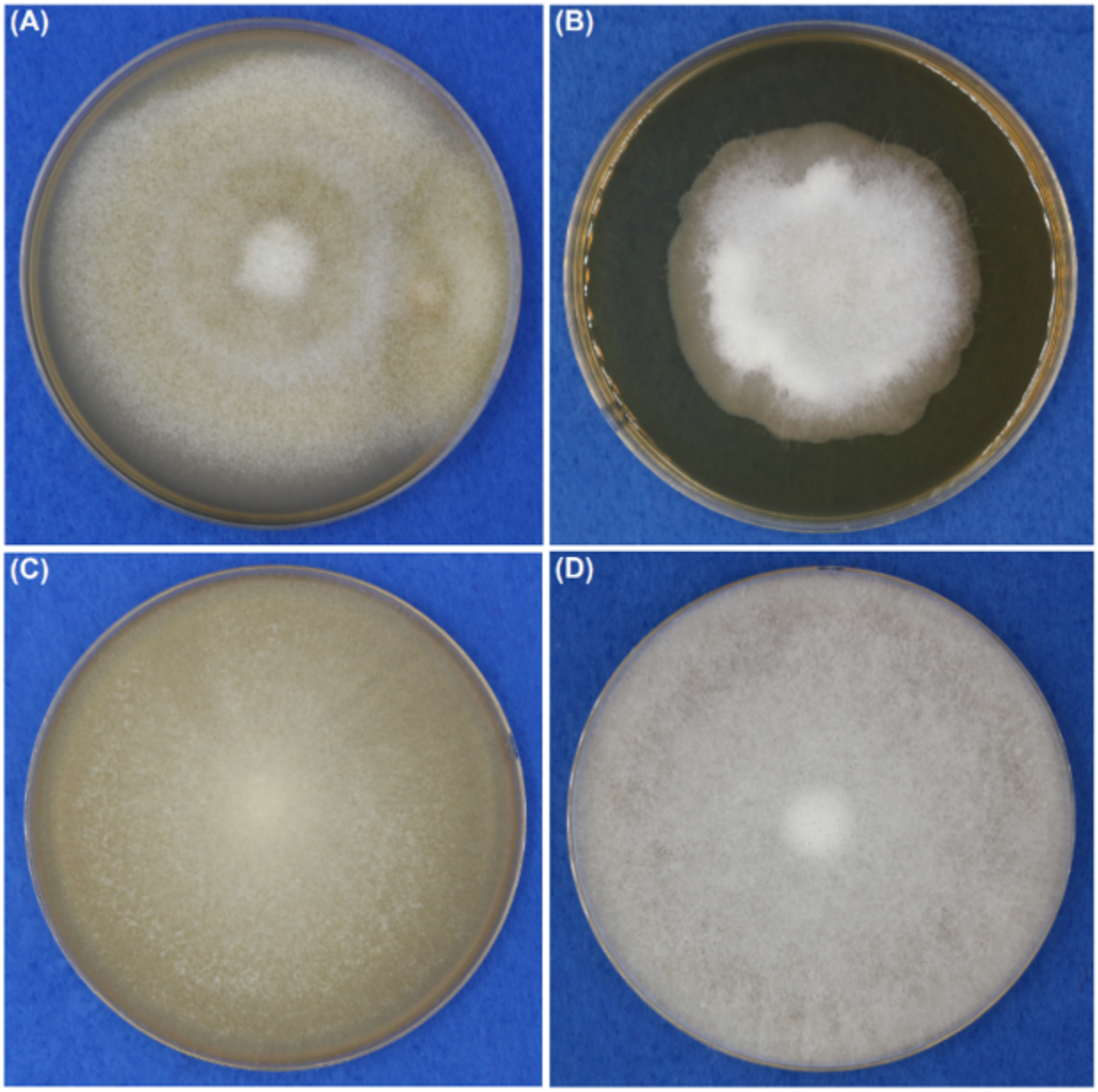



**Figure 9**

Morphology of four *Mucorales* isolates on MEA media after a 4-day culture at 30 °C. (A) *Mucor irregularis* (CMFCCC B50a), (B) *Lichtheimia corymbifera* (CMFCCC B63a), (C) *Mucor hiemalis* (CMFCCC B66h) and (D) *Rhizopus arrhizus* (CMFCCC B81a).

## Supplementary Files

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

- TableS1.xlsx
- TableS2.xlsx
- supplementarymaterial.docx